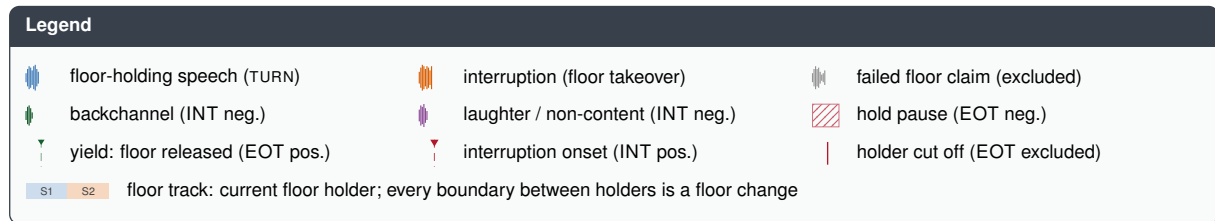


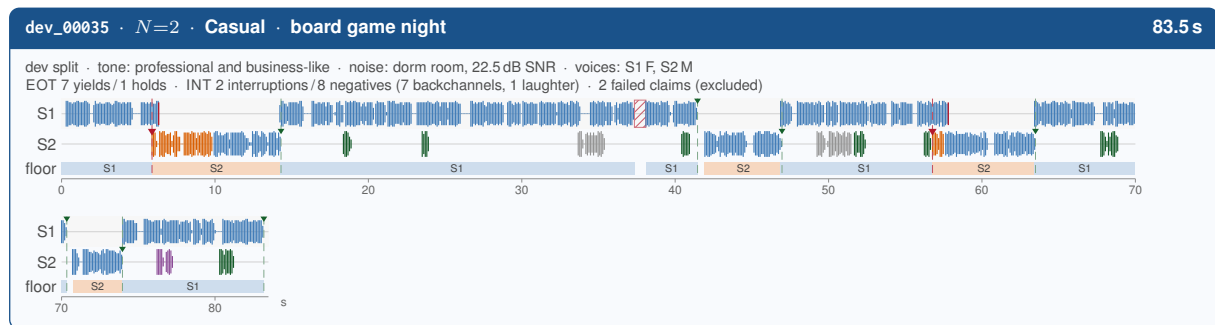
K Data Visualization

This appendix shows every one of the 240 MPTB conversations (48 development, 192 test) as a time-aligned, ground-truth timeline, in the visual vocabulary of Figure 1. Each box is one conversation; boxes are grouped by party count N and ordered by interaction type, and the frame color encodes the interaction type (**casual**, **task-oriented**, **instructional**, **collaborative**, **argumentative**, **narrative**). The box header lists the conversation id, N , type, domain and duration; the line below gives the split, the sampled tone, the noise profile, the speaker voice genders, and the number of scored anchors in the conversation.

Each row covers 70 s on the same horizontal scale in every box, so durations and gaps are directly comparable across conversations. Each speaker has one lane (S1–S N) showing the waveform envelope of that speaker’s clean channel (peak level per 100 ms); benchmark systems receive the noisy channels instead. Waveform color gives the generator label of each utterance, and all markers are derived from the released event logs with the anchor rules of Appendix B.1. Below each box, the full transcript lists every utterance in time order with its realized start and end time (matching the time axis), speaker, ground-truth label in the same color as its waveform, the spoken text as generated, and the floor outcome of each floor-holding segment (yield to the next holder, hold with a mid-turn pause, or cut off by an interruption). No model prediction is shown.



K.1 Conversations with $N = 2$ speakers



Transcript of dev_00035

start–end (s) speaker [ground-truth label] text > outcome

0.3–6.3 S1 [turn] All right, if everyone’s ready, I’ll just go through the basic setup for Power Grid. You’ll each get fifty credits, and the plant cards go— > cut off

5.9–9.9 S2 [interruption] Wait, do I actually get cash or is it just the paper money again? > takes floor from S1

9.9–14.3 S2 [turn] Because I’m still finding the little bills from last session in my backpack. > yields to S1

14.2–20.4 S1 [turn] Sadly, no real money today. That’s just the paper cash, and yes, it does get everywhere if you’re not careful.

18.4–18.9 S2 [backchannel] Right.

20.5–30.0 S1 [turn] So, um, after everyone takes their money, we set up the market with the eight initial power plants. Just remember the left row’s the current market, the right’s future.

23.5–24.0 S2 [backchannel] Uh-huh.

30.1–37.3 S1 [turn] And, for resources, coal and oil go here, uranium and garbage there. The tokens are color coded, so please don’t mix them. > holds, pauses

33.7–35.4 S2 [failed claim] Yeah, mixing is chaos.

38.1–41.5 S1 [turn] —or we’ll end up spending half the night sorting the resource tray again. > yields to S2

40.4–41.0 S2 [backchannel] Noted.

41.9–47.0 S2 [turn] Remind me, what determines turn order again? I always forget after the auctions. > yields to S1

46.9–50.7 S1 [turn] Good question. The player with the most cities goes first, ties broken—

49.2–51.5 S2 [failed claim] By like, highest current power plant?

50.8–52.6 S1 [turn] —by highest-numbered plant, exactly.

51.7–52.4 S2 [backchannel] Makes sense.

52.7–57.8 S1 [turn] So, just to make sure, everybody’s clear on auctions and how you can’t pass and then jump back in, right? > cut off

56.2–56.7 S2 [backchannel] Yeah.

56.8–57.5 S2 [interruption] Oh, but— > takes floor from S1

57.6–63.5 S2 [turn] If someone passes in the first auction and sits out, can they still buy a plant later in the same round? > yields to S1

63.4–70.3 S1 [turn] Yes, but only from the market at face value after all auctions are done, but you can’t take part in the auction phase again. > yields to S2

67.7–68.9 S2 [backchannel] Right, of course.

70.7–74.0 S2 [turn] Sorry, I know these rules are like second nature for you. > yields to S1

74.0–78.5 S1 [turn] Oh, no worries. I’d rather repeat myself than have to pause mid-game for disputes.

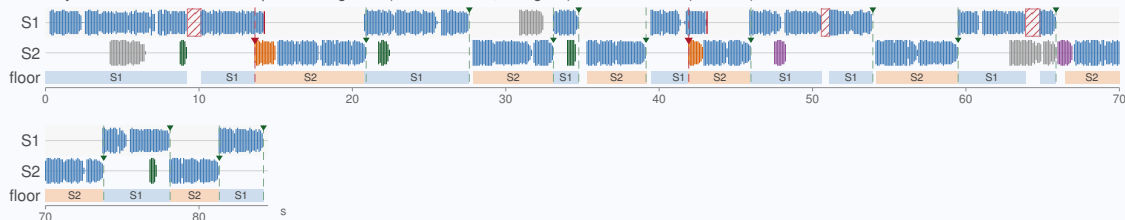
76.2–77.3 S2 [laughter] Haha, true.

78.6–83.2 S1 [turn] All right, I think we’re set. Anything else unclear before we actually start bidding? > yields, end

dev_00046 · $N=2$ · Casual · travel itinerary planning

84.5 s

dev split · tone: curious and inquisitive · noise: cafe, 37.5 dB SNR; alarm beep · voices: S1 M, S2 F
EOT 13 yields/3 holds · INT 2 interruptions/6 negatives (4 backchannels, 2 laughter) · 3 failed claims (excluded)



Transcript of dev_00046

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–7.2 S1 [turn] Alright, so if we're landing in Barcelona at, what, nine in the morning, do you think we should head straight to the hotel, or like try to squeeze something in first?

4.2–6.5 S2 [failed claim] Oh, hotel! I always need to drop my bags—

7.4–9.2 S1 [turn] —yeah, fair, I figured you'd say that. ▷ holds, pauses

8.8–9.2 S2 [backchannel] Mm-hm.

10.1–14.3 S1 [turn] But honestly, if check-in's not till, like, noon, maybe we can grab coffee along Las Ramblas first? That spot with the blue chairs— ▷ cut off

13.7–15.0 S2 [interruption] The one you sent me on Insta? ▷ takes floor from S1

15.2–20.9 S2 [turn] Because that looked so cute, I'm totally down. And I kinda need caffeine after that flight, honestly. ▷ yields to S1

20.8–22.7 S1 [turn] Same, yeah. Jet lag's gonna hit hard.

21.7–22.4 S2 [backchannel] Exactly.

22.8–27.6 S1 [turn] So, after coffee, should we try for the Picasso Museum? Or is that too ambitious on day one? ▷ yields to S2

27.9–31.7 S2 [turn] Uh, honestly, I think it's cool as long as we don't book anything timed—

30.9–32.5 S1 [failed claim] Right, yeah, just kind of wander?

31.8–33.1 S2 [turn] —exactly. No pressure. ▷ yields to S1

33.1–34.8 S1 [turn] Cool, that's my kind of morning. ▷ yields to S2

34.0–34.6 S2 [backchannel] Same here.

35.3–39.1 S2 [turn] What about lunch? Are you still desperate for paella or did you change your mind? ▷ yields to S1

39.5–43.1 S1 [turn] Oh my god, yes, paella forever. There's that place near the water everyone keeps talking about online— ▷ cut off

41.9–42.8 S2 [interruption] La Barceloneta? ▷ takes floor from S1

42.9–46.0 S2 [turn] Yeah, that's meant to be legit. I can try to book if you want. ▷ yields to S1

45.9–50.6 S1 [turn] Haha, you booking stuff already, of course. Yeah, if you can grab us a table, that'd be awesome. ▷ holds, pauses

47.5–48.3 S2 [laughter] Haha, you know me.

51.1–53.9 S1 [turn] How about after lunch, beach? Or will we crash and just nap? ▷ yields to S2

54.1–59.5 S2 [turn] I always say I'll nap but then I just get FOMO. So—maybe beach, maybe siesta, I'm open. ▷ yields to S1

59.5–63.9 S1 [turn] Right, we'll see how we feel. Oh, did you still want to check out that rooftop bar for sunset? ▷ holds, pauses

62.9–65.9 S2 [failed claim] Yes! The one with the live music, hopefully. I'll call ahead—

64.8–65.9 S1 [turn] —of course you will. ▷ yields to S2

66.0–67.0 S2 [laughter] Haha, yeah, you know me.

67.1–73.8 S2 [turn] So just to recap—we land, coffee, Picasso, paella, maybe the beach, rooftop at sunset? Does that sound right? ▷ yields to S1

73.7–78.1 S1 [turn] Honestly, that sounds perfect. Throw in, like, some spontaneous wandering and I'm all in. ▷ yields to S2

76.8–77.2 S2 [backchannel] Mm-hm.

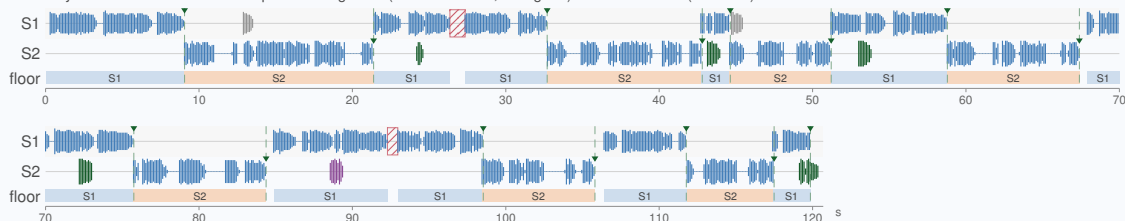
78.1–81.3 S2 [turn] Alright, I'll put this in the shared doc and send you the links tonight. ▷ yields to S1

81.3–84.2 S1 [turn] Awesome, I can't wait! Seriously, this is gonna be so fun. ▷ yields, end

test_00010 · $N=2$ · Casual · family logistics planning

120.7 s

test split · tone: curious and inquisitive · noise: living room tv distant, 33.9 dB SNR; dishes clinking, alarm beep · voices: S1 F, S2 F
EOT 15 yields/2 holds · INT 0 interruptions/6 negatives (5 backchannels, 1 laughter) · 2 failed claims (excluded)



Transcript of test_00010

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.1 S1 [turn] Okay, so if we're gonna do dinner at Mom's Sunday, we actually need to figure out who's driving and when we're leaving, or it's just not gonna happen again. ▷ yields to S2

9.1–15.6 S2 [turn] Yeah, last time was chaos. Like, truly Olympic-level disorganization.

12.9–13.5 S1 [failed claim] I mean, I—

15.7–21.4 S2 [turn] —remember your GPS rerouting us into that parade? Iconic. ▷ yields to S1

21.4–26.4 S1 [turn] —just, that was not my fault, okay? I warned everyone, Google Maps is not psychic. ▷ holds, pauses

24.2–24.6 S2 [backchannel] mm-hm

27.4–32.7 S1 [turn] So, you wanna drive this time, or should we try my car again? It's actually clean, shockingly. ▷ yields to S2

32.7–42.8 S2 [turn] Define 'clean' because last time I think there was an entire gas station's worth of receipts under your seat. ▷ yields to S1

42.7–44.6 **S1** [turn] Ha! Rude. I recycled those. ▷ yields to S2

43.1–44.0 **S2** [backchannel] impressive

44.6–51.2 **S2** [turn] —bring snacks. That was my only real note. If we get stuck, at least chips. ▷ yields to S1

44.7–45.5 **S1** [failed claim] We could just—

51.2–58.8 **S1** [turn] —yes, but what I was trying to say is, we could just leave early. Like actually early, not 'we're leaving but still inside' early. ▷ yields to S2

53.0–53.8 **S2** [backchannel] ambitious

58.8–67.4 **S2** [turn] Okay, but what do you mean by early? Like, before noon, or early for us, which is... what, one-thirty? ▷ yields to S1

67.9–75.8 **S1** [turn] Noon. I think if we aim for noon, we'll realistically get out by twelve forty-five, so that's basically on time in this family. ▷ yields to S2

72.2–73.1 **S2** [backchannel] optimistic

75.7–84.4 **S2** [turn] Wait, do you actually want to drive, or are you trying to pawn it off? Because I can if you really hate it. ▷ yields to S1

84.9–92.3 **S1** [turn] No, I'm fine to drive. I just, I need moral support when—well, when Mom calls like three times while we're in traffic. ▷ holds, pauses

88.6–89.4 **S2** [laughter] haha, yeah

93.0–98.5 **S1** [turn] Maybe you should take her calls so she doesn't yell at me for my driving while I'm already behind the wheel. ▷ yields to S2

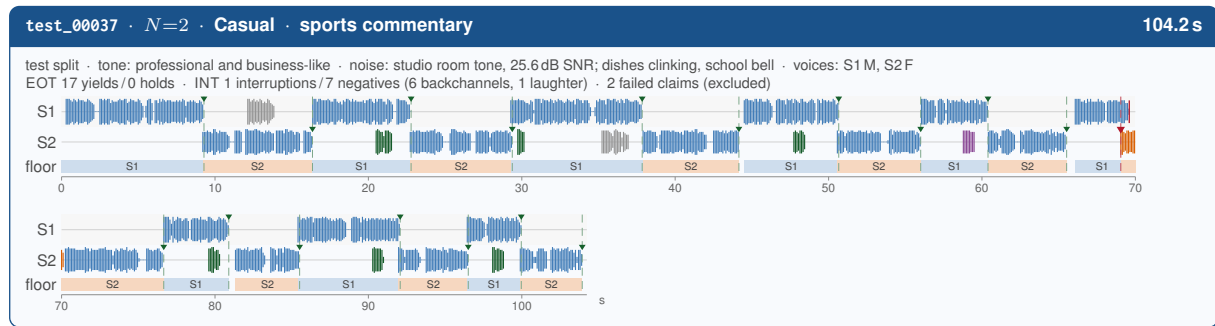
98.5–105.8 **S2** [turn] Bold of you to assume she won't yell at both of us at once. But yeah, that's fair. ▷ yields to S1

106.4–111.8 **S1** [turn] So, it's a plan. Noonish, my car, you're in charge of snacks and running interference. Deal? ▷ yields to S2

111.8–117.5 **S2** [turn] Deal. And if we end up in another parade, dibs on the playlist. ▷ yields to S1

117.4–119.9 **S1** [turn] Deal. But no more ABBA deep cuts, please. ▷ yields, end

119.1–120.4 **S2** [backchannel] zero promises



Transcript of test_00037

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.3 **S1** [turn] Alright, um, welcome back everyone. We're here at halftime for the Tigers versus Mariners game, and uh, I'm still trying to, you know, catch my breath after that last play. ▷ yields to S2

9.2–13.9 **S2** [turn] Can you believe that interception though? I mean, I was almost out of my seat, it was—

12.1–13.9 **S1** [failed claim] Uh, yeah, I honestly thought—

14.0–16.4 **S2** [turn] —like, the way Carter just leaped for it? Unreal. ▷ yields to S1

16.4–22.8 **S1** [turn] —that it was gonna go out of bounds honestly, but Carter's reflexes, like, they, um, caught everyone off guard. That was wild. ▷ yields to S2

20.5–21.6 **S2** [backchannel] Totally, totally.

22.7–29.4 **S2** [turn] Do you think the Mariners even saw that coming? Their offense just, I mean, they froze for like three whole seconds after. ▷ yields to S1

29.3–30.8 **S1** [turn] No, for sure not. I—

29.7–30.2 **S2** [backchannel] Right?

30.9–37.8 **S1** [turn] —think you could see it in their faces right after the play, like, utter disbelief. Their quarterback, uh, he just, like, kept staring at his hands. ▷ yields to S2

35.2–37.0 **S2** [failed claim] Oh my gosh, his eyes were huge.

37.9–44.2 **S2** [turn] I do wonder though, do you think that's gonna shake his confidence for the rest of the game? Or is he the type to bounce back? ▷ yields to S1

44.5–50.7 **S1** [turn] I mean, he's, um, pretty resilient, but something like that, mid-play interception, it can, uh, rattle even the best. ▷ yields to S2

47.7–48.5 **S2** [backchannel] Yeah, yeah.

50.6–56.0 **S2** [turn] But hey, did you notice the Tigers coach during the timeout? He looked like he was doing a little dance out there. ▷ yields to S1

56.1–60.4 **S1** [turn] Haha, I, I did! It was, uh, maybe just adrenaline, but he almost slipped, didn't he? ▷ yields to S2

58.8–59.6 **S2** [laughter] Haha, right?

60.4–65.5 **S2** [turn] Seriously, though, that energy, it's probably why the team's been so aggressive this season. ▷ yields to S1

66.1–69.6 **S1** [turn] Yeah, um, it's definitely infectious. I actually wanted to ask— ▷ cut off

69.0–70.2 **S2** [interruption] Wait, can I just say— ▷ takes floor from S1

70.2–76.7 **S2** [turn] My favorite moment so far was when the camera caught Mitchell just grinning after his block. That was pure gold. ▷ yields to S1

76.7–80.9 **S1** [turn] Oh, right. That smile, I mean, it was almost like he was daring them to try again. ▷ yields to S2

79.6–80.3 **S2** [backchannel] Exactly.

81.3–85.5 **S2** [turn] So what are you expecting in the second half? Any, uh, bold predictions? ▷ yields to S1

85.4–92.1 **S1** [turn] Honestly? I, uh, think it'll come down to who keeps their nerves. If the Tigers push like they have, they might just, you know, lock this one up. ▷ yields to S2

90.3–91.0 **S2** [backchannel] Nice call.

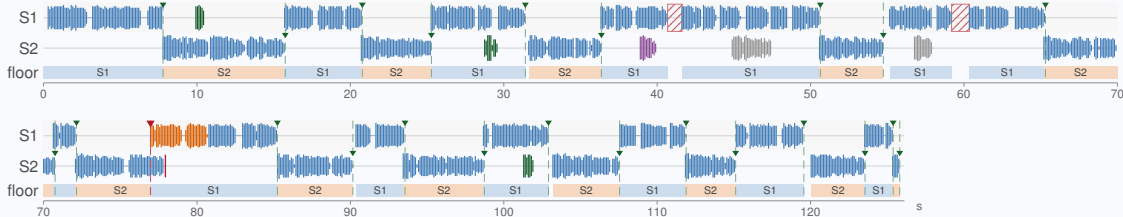
92.0–96.5 **S2** [turn] Alright, we'll see if you're right! I'm betting the Mariners have a surprise up their sleeves, though. ▷ yields to S1

96.4–100.0 **S1** [turn] Haha, maybe, maybe. Either way, it's been, um, a blast so far. ▷ yields to S2

98.1–98.9 **S2** [backchannel] Absolutely.

99.9–103.9 **S2** [turn] Okay folks, stick around, second half's coming up after this quick break. ▷ yields, end

test split · tone: enthusiastic and energetic · noise: studio room tone, 21.1 dB SNR; notification chime, phone ring · voices: S1 F, S2 M
EOT 23 yields / 2 holds · INT 1 interruptions / 4 negatives (3 backchannels, 1 laughter) · 2 failed claims (excluded)



Transcript of test_00056

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–7.8 S1 [turn] Okay, before you say anything, I am dying to know—did you actually finish 'Tomorrow, and Tomorrow, and Tomorrow'? I need your thoughts. ▷ yields to S2

7.8–10.8 S2 [turn] Oh my gosh, yes! I stayed up till like, two in the morning—

9.9–10.5 S1 [backchannel] No way!

10.9–15.8 S2 [turn] —just like inhaling the last hundred pages. I cried at least twice, not even exaggerating. ▷ yields to S1

15.7–20.8 S1 [turn] Yes! The ending kinda destroyed me. I finished and I just sort of, like, stared at my ceiling. ▷ yields to S2

20.7–25.3 S2 [turn] Same! It's—I mean, the characters were so real, I just wanted to shake Sam sometimes, you know? ▷ yields to S1

25.3–31.4 S1 [turn] Right? And Sadie was frustrating but in this way that's like... ugh, I get it—but also, please communicate. ▷ yields to S2

28.8–29.6 S2 [backchannel] Exactly.

31.6–36.4 S2 [turn] So, okay, your turn. Did you finally watch that movie I kept bugging you about, 'The Farewell'? ▷ yields to S1

36.3–40.7 S1 [turn] Oh! Yes, but I had to pause it like, three times because my mom kept calling— ▷ holds, pauses

38.9–39.9 S2 [laughter] Haha, classic!

41.6–47.3 S1 [turn] —but when I finally finished, I was just like, ugly crying in the dark. That movie sneaks up on you.

44.9–47.5 S2 [failed claim] It really does! That ending scene at the airport just—

47.4–50.6 S1 [turn] —oh my god, yes! I literally covered my face with a pillow. ▷ yields to S2

50.6–54.7 S2 [turn] Haha, same. Awkwafina's performance is so understated but so good, right? ▷ yields to S1

55.2–59.2 S1 [turn] Totally! Wait, have you seen her in 'Nora from Queens'? I think you'd love it. ▷ holds, pauses

56.8–58.0 S2 [failed claim] I haven't! Is it on—

60.4–65.3 S1 [turn] —HBO Max! I'll send you my login, honestly. That show is chaos in the best way. ▷ yields to S2

65.2–70.7 S2 [turn] Oh you are a hero, yes please. You know what you'd actually love? Wait, let me recommend a book before I forget— ▷ yields to S1

70.6–72.1 S1 [turn] Hit me! I need something new. ▷ yields to S2

72.1–77.9 S2 [turn] So it's called 'Sea of Tranquility' by Emily St. John Mandel? It's like time travel but not in a headache way, and there's a moon colony— ▷ cut off

77.0–80.7 S1 [interruption] Wait, is that the 'Station Eleven' author? Because I adored that show. ▷ takes floor from S2

80.8–85.2 S1 [turn] No seriously, if it's her, I'm in. I love her vibe—it's always a little dreamy. ▷ yields to S2

85.3–90.2 S2 [turn] Yes, same author! Honestly, I think this one's even weirder but like, in the best possible way. ▷ yields to S1

90.4–93.6 S1 [turn] I'm sold already, adding it to my library list right now. ▷ yields to S2

93.4–98.7 S2 [turn] Haha, perfect. And let me know when you start it so I can fully obsess over every plot twist with you. ▷ yields to S1

98.7–102.9 S1 [turn] Deal! And I'll text you about the first episode of 'Nora from Queens', promise. ▷ yields to S2

101.3–101.9 S2 [backchannel] Can't wait.

103.2–107.6 S2 [turn] Oh, oh! Before I forget—have you read 'The Seven Husbands of Evelyn Hugo'? Because— ▷ yields to S1

107.6–111.9 S1 [turn] Oh my god, yes. I devoured that in, like, two days. Absolute drama. ▷ yields to S2

111.8–115.1 S2 [turn] Right? I just want to live in that old Hollywood world for a second. ▷ yields to S1

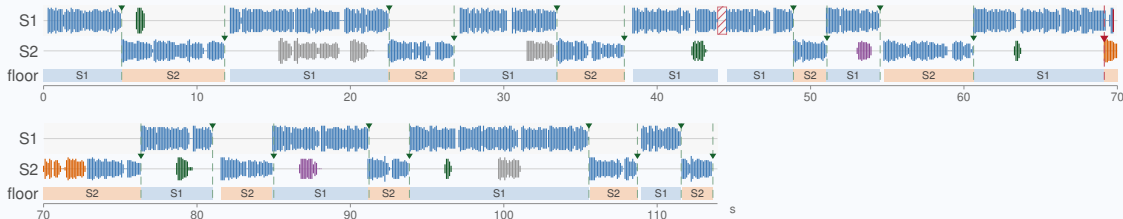
115.2–119.6 S1 [turn] Ugh, same. Okay, we need a whole group chat for this. Rec corner, every week. ▷ yields to S2

120.0–123.5 S2 [turn] I'm in. Okay, I gotta run to the store but text me your thoughts, for real! ▷ yields to S1

123.5–125.4 S1 [turn] Will do! Go get your snacks. Bye!! ▷ yields to S2

125.4–125.8 S2 [turn] Bye! ▷ yields, end

test split · tone: curious and inquisitive · noise: school hallway, 24.8 dB SNR; paper shuffle · voices: S1 F, S2 M
EOT 19 yields / 1 holds · INT 1 interruptions / 7 negatives (5 backchannels, 2 laughter) · 4 failed claims (excluded)



Transcript of test_00070

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–5.1 S1 [turn] Okay, so, um, should we start with the bio review or do you wanna look at the old chem test first? ▷ yields to S2

5.1–7.1 S2 [turn] Bio's good, yeah, I still don't get like—

6.1–6.6 S1 [backchannel] mm-hm

7.2–11.8 S2 [turn] —the difference between facilitated diffusion and active transport? It always throws me off. ▷ yields to S1

12.2–16.4 S1 [turn] Yeah, fair. Facilitated's like, no energy, just a channel or something helping, right?

15.4–19.4 S2 [failed claim] But isn't there, like, ATP involved in both or am I making that up—

16.4–22.5 S1 [turn] —No, it's only in active transport that you use ATP. Facilitated's just, uh, things moving down their gradient? ▷ yields to S2

20.0–21.6 S2 [failed claim] Oh right, yeah yeah.

22.5–26.8 S2 [turn] So, that sodium-potassium pump thing, that's active? Or is it something else? ▷ yields to S1

27.2–33.5 S1 [turn] Yeah, exactly, that's like the classic active transport example. It pumps against the gradient so, uh, ATP needed. ▷ yields to S2

31.5–33.3 **S2** [failed claim] Okay, okay, yeah, makes sense.

33.5–37.9 **S2** [turn] What about osmosis, is that just, like, another word for diffusion but for water? ▷ yields to S1

38.4–43.9 **S1** [turn] Pretty much! It's water moving through a membrane, but only if it's semi-permeable. And still no energy needed. ▷ holds, pauses

42.3–43.3 **S2** [backchannel] Ah, gotchu.

44.5–48.9 **S1** [turn] Did she say, uh, they're gonna ask about plasmolysis too? I heard someone mention that. ▷ yields to S2

48.9–51.1 **S2** [turn] Yeah, dude, that's when the cell shrivels, right? ▷ yields to S1

51.0–54.5 **S1** [turn] Haha, yeah like, if it loses water in a super salty solution. ▷ yields to S2

53.0–54.0 **S2** [laughter] haha right

54.8–60.6 **S2** [turn] Okay so, switching gears, do you remember what she said about enzyme denaturation? I zoned out, I admit. ▷ yields to S1

60.6–66.3 **S1** [turn] Oh, um, yeah, that's when the enzyme kinda, like, unfolds. High heat, wrong pH, it just, uh, loses its shape.

63.3–63.8 **S2** [backchannel] mm-hm

66.4–69.7 **S1** [turn] And if it loses shape, it can't bind to the substrate anymore. So, basically, it stops working. ▷ cut off

69.1–72.7 **S2** [interruption] But sometimes it can fold back, can't it? Or is that just, like, rare? ▷ takes floor from S1

72.9–76.4 **S2** [turn] I remember her saying sometimes they renature, but not always. ▷ yields to S1

76.4–81.0 **S1** [turn] Yeah, I think it's, like, only if the denaturation wasn't too bad? But usually it's permanent. ▷ yields to S2

78.7–79.8 **S2** [backchannel] Right, yeah.

81.6–85.0 **S2** [turn] Honestly, I wish there was a, you know, cheat code for all these terms. ▷ yields to S1

84.9–88.0 **S1** [turn] Haha, right? If only Quizlet could just beam it into our brains.

86.7–88.2 **S2** [laughter] haha exactly

88.1–91.2 **S1** [turn] Wait, did you do the photosynthesis practice problems or nah? ▷ yields to S2

91.2–93.8 **S2** [turn] Not yet, no, do I need to? Were they actually hard? ▷ yields to S1

93.9–96.9 **S1** [turn] Only a couple are tricky, but I can walk you through them if you want.

96.2–96.6 **S2** [backchannel] mm-hm

97.1–102.8 **S1** [turn] So, basically, there's one about light-dependent versus light-independent reactions, and it's super easy to mix them up—

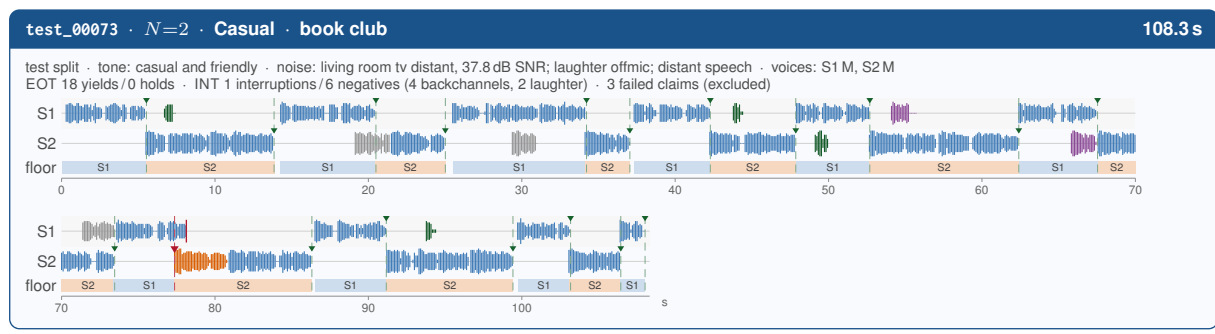
99.7–101.2 **S2** [failed claim] Wait, wait, let me guess—

102.9–105.5 **S1** [turn] —Yeah! What's the product of the light-dependent reaction? ▷ yields to S2

105.6–108.7 **S2** [turn] Uh, ATP and NADPH, right? Oxygen too? ▷ yields to S1

109.0–111.6 **S1** [turn] Perfect, yes, exactly! So you totally got this. ▷ yields to S2

111.6–113.6 **S2** [turn] Sweet. Thanks, that's reassuring. ▷ yields, end



Transcript of test_00073

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–5.5 **S1** [turn] So, uh, did you actually finish the book this time, or did you get distracted by, you know, football again? ▷ yields to S2

5.5–9.7 **S2** [turn] Hey! I totally finished it! Well, okay, I skimmed like the last two chapters, but—

6.7–7.4 **S1** [backchannel] Knew it.

9.8–13.9 **S2** [turn] —I did get the main point. Like, that twist with the neighbor, I didn't see that coming at all. ▷ yields to S1

14.3–20.5 **S1** [turn] Yeah, I mean, the neighbor being the one leaving the weird notes—honestly, I was suspecting the ex-boyfriend for most of it. ▷ yields to S2

21.1–21.4 **S2** [failed claim] Same, same! The author totally set that up.

21.5–25.0 **S2** [turn] But, okay, did you catch all the movie references? There were so many! ▷ yields to S1

25.5–31.1 **S1** [turn] Honestly, half of 'em went over my head. I've, uh, never seen that classic horror series she kept mentioning.

29.4–31.0 **S2** [failed claim] Wait, what? You've never seen it?

31.2–34.2 **S1** [turn] —Nope. I know, I know. I'm culturally deprived, apparently. ▷ yields to S2

34.1–37.0 **S2** [turn] Haha, get out! We're watching it next week. That's a crime. ▷ yields to S1

37.3–42.3 **S1** [turn] We'll see. I might fall asleep. Anyway, what'd you think of the ending? Felt a little, uh, rushed to me. ▷ yields to S2

42.3–47.9 **S2** [turn] Yeah, totally. I was like, wait, that's it? I wanted one more chapter with everyone, you know, actually talking it out. ▷ yields to S1

43.8–44.5 **S1** [backchannel] Right.

47.9–52.7 **S1** [turn] Honestly, if they just had, like, a single conversation, half the drama would've vanished. ▷ yields to S2

49.1–50.0 **S2** [backchannel] Exactly!

52.7–57.0 **S2** [turn] But that's, like, every mystery novel ever. People refusing to talk to each other.

54.1–55.7 **S1** [laughter] Ha, you got a point.

57.2–62.4 **S2** [turn] Okay, but here's what I really wanna know: did you relate to the main character at all, or was it just me? ▷ yields to S1

62.4–67.5 **S1** [turn] Uh, not gonna lie, she reminded me of my cousin. Like, the way she overthinks literally everything. ▷ yields to S2

65.8–67.4 **S2** [laughter] Haha right, that's so true.

67.6–73.5 **S2** [turn] I loved the bit where she, uh, tries to make that midnight pie and just totally burns it. That felt so real. ▷ yields to S1

71.4–73.5 **S1** [failed claim] Yeah, that's, like, peak disaster energy.

73.6–78.1 **S1** [turn] Though I think my favorite scene was the old bookstore. I mean, wish there were still places like that. ▷ cut off

77.4–80.8 **S2** [interruption] Oh! I actually know a place like that on Main. We should go sometime. ▷ takes floor from S1

80.9–86.2 **S2** [turn] Seriously, you'd love it. Smells like old paper and coffee. It's got the creaky floors and everything. ▷ yields to S1

86.5–91.2 **S1** [turn] Heh, alright, I'm in. As long as you don't make me, uh, read poetry out loud or something. ▷ yields to S2

91.2–96.5 **S2** [turn] Promise. No public poetry. Just browsing and, I dunno, maybe judging each other's taste in covers.

93.8–94.4 **S1** [backchannel] Deal.

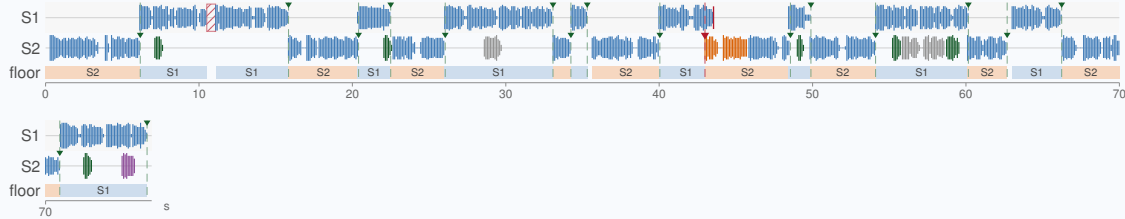
96.7–99.4 **S2** [turn] Oh, before I forget—next book, your pick or mine? ▷ yields to S1

99.7–103.2 **S1** [turn] Your turn, but if you pick another 500 pager, I'm out, man. ▷ yields to S2

103.1–106.5 **S2** [turn] Haha, okay, okay. I swear, this one's, like, two hundred tops. ▷ yields to S1

106.4–108.0 **S1** [turn] I'll believe it when I see it. ▷ yields, end

test split · tone: casual and friendly · noise: school hallway, 22.4 dB SNR; paper shuffle · voices: S1 M, S2 M
 EOT 17 yields / 1 holds · INT 1 interruptions / 7 negatives (6 backchannels, 1 laughter) · 2 failed claims (excluded)

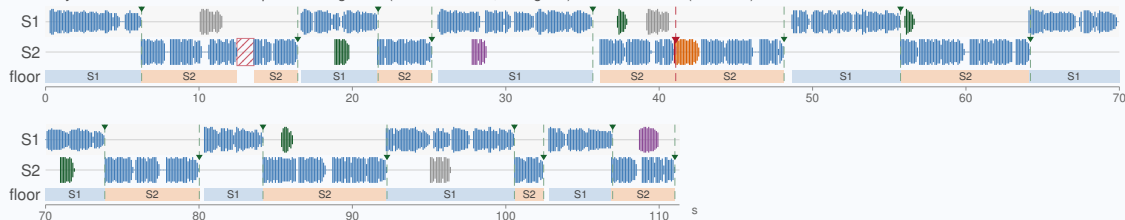


Transcript of test_00100

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.2 S2 [turn] Alright, so do you wanna figure out a grocery list for the week, or are we just kind of winging it again? ▷ yields to S1
 6.1–10.5 S1 [turn] Heh, man, winging it hasn't exactly kept the fridge from looking like a desert, y'know? ▷ holds, pauses
 7.1–7.6 S2 [backchannel] true
 11.1–15.8 S1 [turn] But... I'm down to try actually planning for once. What do you think, tacos on Tuesday again? ▷ yields to S2
 15.8–20.4 S2 [turn] Yes please. Honestly, if we just rotate tacos and pasta, I'm good. ▷ yields to S1
 20.3–22.5 S1 [turn] Haha, well, I do love carbs, man. ▷ yields to S2
 22.1–22.6 S2 [backchannel] same
 22.7–26.1 S2 [turn] Wait, what about breakfast stuff? We're, uh, out of eggs, right? ▷ yields to S1
 26.0–29.4 S1 [turn] Oh yeah, totally out. Also we're dangerously low on coffee.
 28.6–29.7 S2 [failed claim] No, that can't happen—
 29.6–33.1 S1 [turn] —yeah, I know! Crisis. So I'll put coffee and eggs at the top, got it. ▷ yields to S2
 33.1–34.2 S2 [turn] Thanks, appreciate you. ▷ yields to S1
 34.3–35.3 S1 [turn] No prob, man. ▷ yields to S2
 35.6–40.0 S2 [turn] So... do you wanna split up the shopping or just, like, whoever gets to it first? ▷ yields to S1
 40.0–43.5 S1 [turn] I can go after work Monday, if that's cool. Unless you wanna tag along? ▷ cut off
 43.0–45.8 S2 [interruption] Uh, actually, I think I have that lab meeting— ▷ takes floor from S1
 45.9–48.6 S2 [turn] Yeah, never mind, Monday's rough for me. All you. ▷ yields to S1
 48.4–49.9 S1 [turn] Alright, all good. ▷ yields to S2
 49.0–49.4 S2 [backchannel] mm-hm
 49.8–54.1 S2 [turn] Can you check if we're out of TP, too? I swear we go through it so fast. ▷ yields to S1
 54.1–56.9 S1 [turn] Already checked, we've got like two rolls. I'll grab more.
 55.2–55.8 S2 [backchannel] legend
 55.9–58.7 S2 [failed claim] Oh, what about cleaning? We skipped last weekend—
 57.0–60.2 S1 [turn] —oof, yeah, my bad! The bathroom's looking... not its best. ▷ yields to S2
 58.7–59.6 S2 [backchannel] no worries
 60.1–62.7 S2 [turn] I can do kitchen if you wanna take the bathroom? ▷ yields to S1
 63.0–66.2 S1 [turn] Deal. Should we, like, actually set a day? Or just when we remember? ▷ yields to S2
 66.3–70.9 S2 [turn] Sunday morning? Get it over with early, then it's just coffee and laziness. ▷ yields to S1
 71.0–73.8 S1 [turn] Haha, that's the dream. Alright, Sunday it is.
 72.5–73.1 S2 [backchannel] Nice.
 74.0–76.6 S1 [turn] You know, we might actually have our act together this week. ▷ yields, end
 75.0–75.9 S2 [laughter] haha we'll see

test split · tone: curious and inquisitive · noise: dorm room, 21.5 dB SNR; alarm beep, laughter offmic · voices: S1 F, S2 M
 EOT 16 yields / 1 holds · INT 1 interruptions / 7 negatives (5 backchannels, 2 laughter) · 3 failed claims (excluded)



Transcript of test_00173

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.3 S1 [turn] Alright, so, I figured I'd just ping you here about the sign-up sheets for Saturday's cleanup. Did you get my email? ▷ yields to S2
 6.2–12.5 S2 [turn] Yeah I saw it this morning, um, but I haven't actually opened the document yet. Sorry, crazy start to the day. ▷ holds, pauses
 10.1–11.6 S1 [failed claim] No worries, it's just—
 13.6–16.4 S2 [turn] —I will! I promise. Was there something urgent about it? ▷ yields to S1
 16.7–21.7 S1 [turn] It's not urgent, but Jenny wanted a headcount for, like, supplies and snacks by tomorrow I think. ▷ yields to S2
 18.9–19.8 S2 [backchannel] Oh, right, right.
 21.6–25.2 S2 [turn] Did we ever figure out if she got that confirmation from the park office? ▷ yields to S1
 25.6–31.8 S1 [turn] She copied me on something yesterday but it was super vague. Like—it just said, 'Should be fine.' Not super reassuring.
 27.8–28.8 S2 [laughter] Haha, classic.
 32.0–35.7 S1 [turn] Exactly! So uh, I guess worst case, we just show up and improvise? ▷ yields to S2
 36.1–37.8 S2 [turn] I mean, wouldn't be the first time.
 37.3–37.9 S1 [backchannel] Mhm.

37.9–41.0 **S2** [turn] So what's the plan if it rains though? Did you see the forecast? > cut off

39.2–40.7 **S1** [failed claim] Yeah, I checked, it's like—

41.1–42.6 **S2** [interruption] Fifty percent or something? > takes floor from S1

42.8–48.1 **S2** [turn] Because I was telling my sister, if it's like dumping, I'm just not—I'm not dragging people out there. > yields to S1

48.7–55.7 **S1** [turn] Totally fair, yeah. Jenny said we'd just reschedule if it's actually storming, so, backup plan is basically 'try again next week.' > yields to S2

55.7–56.9 **S2** [turn] Alright. That's easy.

56.0–56.6 **S1** [backchannel] Mhm.

57.0–64.2 **S2** [turn] Is there a particular spot in the park you want to tackle first? Because last time folks just wandered off everywhere, it got a little messy. > yields to S1

64.1–73.9 **S1** [turn] Oh, yeah, uh, I marked a few zones on the signup sheet, actually. Like, playground, creek trail, picnic area. I was hoping to assign people, but only if that's not too, like, controlling? > yields to S2

71.0–71.9 **S2** [backchannel] No, that's good.

73.9–80.0 **S2** [turn] People honestly like having a spot, so they're not just waiting around. Maybe even print out a little map or something? > yields to S1

80.4–84.2 **S1** [turn] That's a good idea. Mm, I could just, like, sketch one out and bring a few copies? > yields to S2

84.2–86.4 **S2** [turn] Yeah! Or even text a photo if people forget.

85.4–86.2 **S1** [backchannel] Smart.

86.5–92.3 **S2** [turn] Oh, wait, do we have enough gloves left? Because someone took like five pairs last month and I have no idea who. > yields to S1

92.2–96.4 **S1** [turn] Ugh, no clue. I think I saw a stray bag in my trunk though. I'll check and text you—

95.1–96.4 **S2** [failed claim] Thanks. You're a lifesaver.

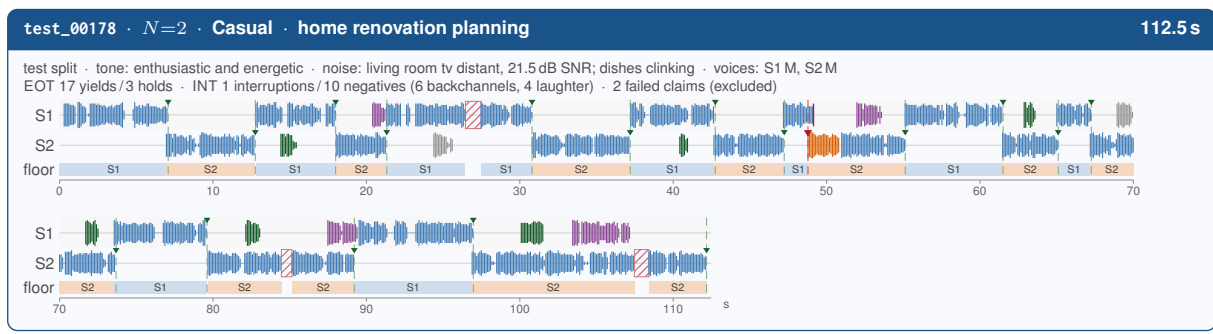
96.5–100.6 **S1** [turn] —if I find them tonight. Otherwise, I'll stop at the supply store Friday. > yields to S2

100.6–102.5 **S2** [turn] That's awesome. Seriously, thank you. > yields to S1

102.8–107.0 **S1** [turn] Yeah, no problem. Just ping me tomorrow if you've got any other ideas, okay? > yields to S2

106.9–111.0 **S2** [turn] Will do. I'm excited, honestly. It's always kinda chaotic but fun, you know? > yields, end

108.7–110.0 **S1** [laughter] Haha, totally.



Transcript of test_00178

start—end (s) speaker [ground-truth label] text > outcome

0.3–7.1 **S1** [turn] Okay, so, honestly, I think the living room is where we should start. That's the space everybody sees, you know? > yields to S2

7.0–12.8 **S2** [turn] Yeah, yeah, first impressions, right. I was staring at that wall where the paint's all chipped and thought, man, if we just— > yields to S1

12.8–18.0 **S1** [turn] Oh, for sure, it's so obvious! That, like, weird green from the seventies? Has to go. > yields to S2

14.4–15.5 **S2** [backchannel] God, yes.

18.0–21.3 **S2** [turn] I can't believe anyone ever thought avocado green was stylish. > yields to S1

20.4–21.2 **S1** [laughter] Haha, right?

21.3–26.5 **S1** [turn] So, paint first, or—are you still thinking about knocking out that little divider wall? > holds, pauses

24.4–25.6 **S2** [failed claim] Well, hang on—

27.5–30.8 **S1** [turn] —I mean, if it's not load-bearing, it'd open everything up. > yields to S2

30.8–37.2 **S2** [turn] Exactly. So, I was talking to Dave? He thinks we can demo it pretty easily but suggests we double check the wiring first. > yields to S1

37.2–42.7 **S1** [turn] Smart. Yeah, we do not wanna accidentally cut into, like, a giant mess of wires. > yields to S2

40.4–41.0 **S2** [backchannel] Mm-hm.

42.7–47.2 **S2** [turn] So, maybe, step one: call in an electrician before I start swinging the sledgehammer. > yields to S1

47.2–49.1 **S1** [turn] Probably a safer bet than you just— > cut off

48.8–50.9 **S2** [interruption] Oh come on, where's your sense of adventure? > takes floor from S1

51.0–55.1 **S2** [turn] I mean, worst case, the power goes out for a few hours. Not the end of the world. > yields to S1

52.0–53.6 **S1** [laughter] Haha, famous last words.

55.1–61.5 **S1** [turn] Okay but, after the wall, what about the floors? That old carpet... I mean, it's gotta be, what, '92? > yields to S2

61.5–65.1 **S2** [turn] Oh, easily. There's probably enough dust in there to start a new planet. > yields to S1

62.9–63.6 **S1** [backchannel] Ugh, yeah.

65.0–67.3 **S1** [turn] Have you decided hardwood or, like, laminate? > yields to S2

67.2–70.3 **S2** [turn] So, I keep going back and forth. Hardwood looks amazing but—

68.9–70.0 **S1** [failed claim] But the price, right?

70.4–73.7 **S2** [turn] —yeah, exactly. And with the kids? Laminate's just... easier. > yields to S1

71.7–72.5 **S1** [backchannel] Totally get it.

73.6–79.6 **S1** [turn] Oh, have you seen those samples at the hardware store? Some of the laminate looks so real now, it's wild. > yields to S2

79.7–84.5 **S2** [turn] Yeah! I almost grabbed a handful to show you, but then Mason started climbing the sample display— > holds, pauses

82.1–83.1 **S1** [backchannel] Classic Mason.

85.2–89.2 **S2** [turn] He just, like, wanted to swing from the shelves. The poor sales guy looked terrified. > yields to S1

87.5–89.5 **S1** [laughter] Haha, oh man, I wish I'd seen that.

89.5–97.0 **S1** [turn] Alright, paint, wall, floor—what about lighting? You ever think about recessed lights instead of that big chandelier thing? > yields to S2

96.9–101.6 **S2** [turn] Oh! You read my mind. I literally just found a sale on LED recessed kits yesterday.

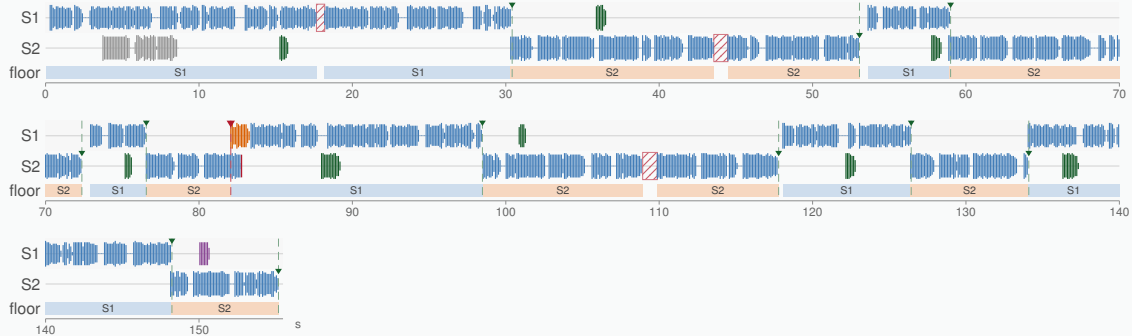
100.1–101.6 **S1** [backchannel] Seriously? That's perfect.

101.7–107.5 **S2** [turn] We could swap them all out in an afternoon, honestly. Even with my, uh, questionable wiring skills. > holds, pauses

103.5–107.2 **S1** [laughter] Haha, as long as you don't electrocute yourself twice in one month.

108.4–112.2 **S2** [turn] No promises, man. But hey, that's what you get when you let me run wild. > yields, end

dev split · tone: enthusiastic and energetic · noise: studio room tone, 37.5 dB SNR · voices: S1 F, S2 M
EOT 11 yields / 3 holds · INT 1 interruptions / 9 negatives (8 backchannels, 1 laughter) · 1 failed claims (excluded)

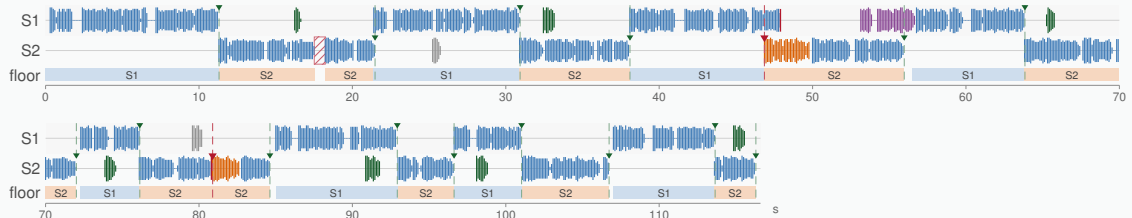


Transcript of dev_00030

start-end (s) speaker [ground-truth label] text ▷ outcome

- 0.3–10.1 **S1** [turn] Welcome back to the Roundtable, everyone! Today we're diving into the new urban transport initiative announced yesterday. There's a lot to unpack, so let's jump right in—
- 3.7–8.7 **S2** [failed claim] Yeah, before we get into the details, can I just say this rollout caught me by surprise—
- 10.2–17.7 **S1** [turn] —right? The city council didn't tease it at all. But I want to start with the goals, actually, since that's where most listeners have questions. ▷ holds, pauses
- 15.3–15.8 **S2** [backchannel] mm-hm
- 18.2–30.4 **S1** [turn] So the core idea is to reduce downtown congestion by incentivizing public transit. They're adding those new electric buses and also expanding the bike-share. But, um, Russ, you've crunched the numbers— ▷ yields to S2
- 30.3–43.6 **S2** [turn] Oh, yeah, absolutely. The initial budget estimate is just over 200 million dollars, but that assumes ridership increases by, get this, 30 percent in the first year. I'm not sure that's realistic— ▷ holds, pauses
- 35.9–36.6 **S1** [backchannel] No way
- 44.5–53.1 **S2** [turn] —given current trends, anyway. Part of the funding depends on how quickly they get the infrastructure up, so delays could blow up the costs. ▷ yields to S1
- 53.6–59.0 **S1** [turn] Exactly. And so many projects have run into delays because of supply chain issues lately. ▷ yields to S2
- 57.8–58.4 **S2** [backchannel] right
- 58.9–72.4 **S2** [turn] But you know, one thing that stood out to me was their emphasis on equity—making sure the new routes actually serve neighborhoods that need better options the most. That's, uh, an area a lot of plans just gloss over. ▷ yields to S1
- 72.9–76.6 **S1** [turn] Yeah, totally. And I thought their map, did you get a chance to look at it? ▷ yields to S2
- 75.2–75.7 **S2** [backchannel] uh-huh
- 76.6–82.8 **S2** [turn] They actually included North Park, which never gets included! I mean, okay, it still only gets one new stop, but it's something. But— ▷ cut off
- 82.1–83.3 **S1** [interruption] Wait, let me jump in— ▷ takes floor from S2
- 83.4–92.3 **S1** [turn] I actually spoke to a resident from North Park this morning, and she was skeptical. Apparently the schedules don't match the early shifts a lot of folks work there.
- 88.0–89.2 **S2** [backchannel] Oh interesting
- 92.5–98.5 **S1** [turn] So, inclusion is great on paper, but there's still a disconnect if the service doesn't fit real life, you know? ▷ yields to S2
- 98.5–108.9 **S2** [turn] That's a huge point. Honestly, the way agencies measure success doesn't always map to actual usability. Like, they'll count stops, but not if anyone can use them. ▷ holds, pauses
- 100.9–101.4 **S1** [backchannel] yeah
- 109.9–117.8 **S2** [turn] Also, did you notice the new fare structure? They're piloting an income-based discount. That could really make a difference if it works. ▷ yields to S1
- 118.1–126.4 **S1** [turn] Absolutely, especially if it's paired with that mobile payment rollout. Although, I'm a little worried about access for folks who don't use smartphones. ▷ yields to S2
- 122.2–122.8 **S2** [backchannel] right
- 126.4–134.1 **S2** [turn] I was just about to say that. There's always this tech optimism, but it leaves out people without the latest device. Uh— ▷ yields to S1
- 134.1–141.0 **S1** [turn] Yes, and the city's plan mentioned kiosks at hub stations, but nothing about support for, say, non-English speakers.
- 136.3–137.3 **S2** [backchannel] Exactly
- 141.1–148.2 **S1** [turn] So I think, bottom line, big ambitions—maybe even the right ambitions—but like always, the challenges are in the details. ▷ yields to S2
- 148.2–155.2 **S2** [turn] Couldn't have put it better. I'm genuinely excited to see what happens, but, you know, optimism with a grain of salt. ▷ yields, end
- 150.1–150.7 **S1** [laughter] haha yeah

dev split · tone: curious and inquisitive · noise: office hum, 33.0 dB SNR · voices: S1 F, S2 M
EOT 15 yields / 1 holds · INT 2 interruptions / 8 negatives (7 backchannels, 1 laughter) · 2 failed claims (excluded)



Transcript of dev_00031

start-end (s) speaker [ground-truth label] text ▷ outcome

- 0.3–11.3 **S1** [turn] Alright, let's get started. Our main goal today is sorting out roles for the park clean-up on Saturday, right? I just want to make sure we're both on the same page before diving into specifics. ▷ yields to S2
- 11.3–17.5 **S2** [turn] Yeah, exactly, I was thinking maybe we could rotate jobs so no one gets stuck just picking up trash the whole time— ▷ holds, pauses
- 16.2–16.7 **S1** [backchannel] mm-hm
- 18.2–21.5 **S2** [turn] —unless people actually want to, but, you know, mix it up a little. ▷ yields to S1
- 21.4–27.2 **S1** [turn] That makes sense, but do you think we need to formally assign rotations or just let folks swap on their own?

25.2–25.8 **S2** [failed claim] Well—

27.3–30.9 **S1** [turn] —I mean, I guess it depends on how many volunteers we actually get. ▷ yields to S2

30.9–38.1 **S2** [turn] Right, that's what I was gonna say! If there's only, say, ten of us, we'll probably need assignments. But if twenty show up, maybe it's less rigid. ▷ yields to S1

32.4–33.2 **S1** [backchannel] Exactly

38.1–47.9 **S1** [turn] Okay, so I'll send out an RSVP reminder tonight. But we do need someone at registration, someone to hand out gloves and bags, and at least two people for the stream area, right? ▷ cut off

46.9–49.9 **S2** [interruption] Uh, yeah, and we can't forget about water stations— ▷ takes floor from S1

50.0–56.0 **S2** [turn] —last year people kept asking where to get water. Maybe we double up and have a roaming person check on everyone? ▷ yields to S1

53.1–56.7 **S1** [laughter] Haha, yeah, the water situation got a little desperate by noon.

56.8–63.8 **S1** [turn] I can do the first shift at registration, if that helps. But honestly, I do want to actually pick up trash this time too. ▷ yields to S2

63.8–67.5 **S2** [turn] No problem, I can cover bags and gloves for the first round then—

65.3–65.8 **S1** [backchannel] thanks

67.6–72.0 **S2** [turn] —and maybe after an hour we just swap? Also, who's bringing the sign-up sheets? ▷ yields to S1

72.3–76.1 **S1** [turn] Oh, good catch. I almost forgot. I can print those out tomorrow at work. ▷ yields to S2

73.9–74.6 **S2** [backchannel] Perfect.

76.1–80.8 **S2** [turn] For the stream area, do we need any extra gear? I remember someone lost a boot last time. ▷ cut off

79.6–80.3 **S1** [failed claim] Uh, yes—

80.9–82.7 **S2** [interruption] Haha, it was Trevor, right? ▷ takes floor from S1

82.8–84.6 **S2** [turn] He never heard the end of that. ▷ yields to S1

85.0–92.9 **S1** [turn] Right, exactly. I'll email everyone tonight to bring old boots if they have them. If not, I think I still have some extras in storage. ▷ yields to S2

90.9–91.8 **S2** [backchannel] Oh great.

93.0–96.6 **S2** [turn] Anything else? Food? Or is that overkill for just a half day? ▷ yields to S1

96.6–101.0 **S1** [turn] Hm, maybe just snacks? Nothing fancy, but granola bars and fruit couldn't hurt. ▷ yields to S2

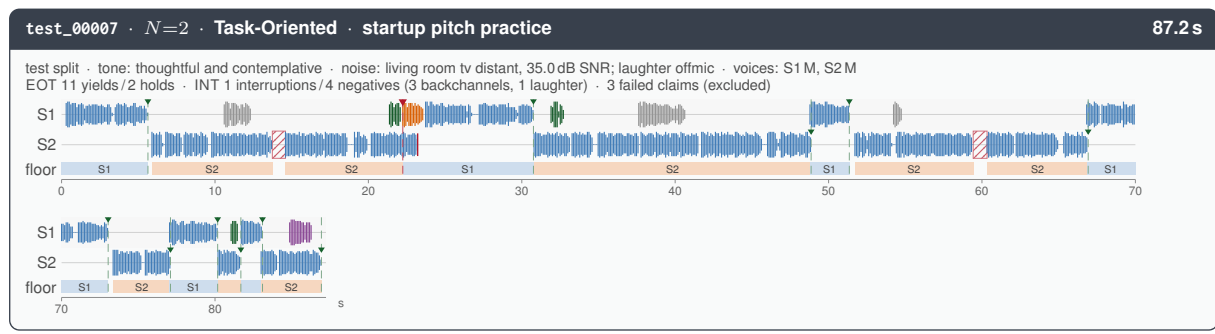
98.1–98.9 **S2** [backchannel] Agreed.

101.1–106.7 **S2** [turn] Alright, so RSVP reminder, print sign-ups, extra boots, water stations, and snacks. That it? ▷ yields to S1

107.0–113.6 **S1** [turn] Yes, I think that covers the main stuff. I'll email everyone tonight and we can revisit on Thursday if things change. ▷ yields to S2

113.6–116.3 **S2** [turn] Cool, thanks for organizing everything again. ▷ yields, end

114.9–115.6 **S1** [backchannel] Of course!



Transcript of test_00007

0.3–5.6 **S1** [turn] Alright, Jay, let's do a run-through just like Demo Day, yeah? Whenever you're ready, just start with your opener. ▷ yields to S2

5.9–13.7 **S2** [turn] Um, okay. Hi, my name is Jay, and I'm the co-founder of Fresco, where we help small restaurants reduce food waste using smart inventory tracking. ▷ holds, pauses

10.6–12.3 **S1** [failed claim] Nice. Good energy so far.

14.6–23.2 **S2** [turn] Um—each year, restaurants in the US waste over thirty billion pounds of food. That costs them—uh—over \$160 billion and hurts the environment, too. ▷ cut off

21.4–22.2 **S1** [backchannel] Right, yeah—

22.2–23.6 **S1** [interruption] Wait, can I jump in— ▷ takes floor from S2

23.7–30.8 **S1** [turn] That stat is really striking, but it feels a little heavy up front. Maybe lead with a short story about a customer, then hit the data? ▷ yields to S2

30.8–33.2 **S2** [turn] Oh, um, like—start with a, a restaurant owner?

31.9–32.8 **S1** [backchannel] Exactly.

33.3–41.6 **S2** [turn] Okay. Um, let me try. Last month, Maria, who owns Little Spoon Café, called us because she was throwing out trays of unsold sandwiches every week.

37.6–40.7 **S1** [failed claim] Yeah, perfect! I can totally picture Maria. Keep going.

41.7–48.9 **S2** [turn] With our system, Maria started tracking which sandwiches sold on which days, and within a month, her waste was down by forty percent. ▷ yields to S1

48.8–51.4 **S1** [turn] Love it. Maybe just tighten the wording, but the story works. ▷ yields to S2

51.7–59.4 **S2** [turn] Alright, um, our product uses a simple app, so restaurant owners can scan inventory daily and get automatic alerts before things expire. ▷ holds, pauses

54.2–54.8 **S1** [failed claim] Wait—

60.3–66.9 **S2** [turn] —So, uh, it saves time and money, and also helps the environment by cutting down on waste. Sorry, you wanted to say something? ▷ yields to S1

66.8–73.0 **S1** [turn] No, you covered it! I was just gonna say, maybe highlight that you're already piloting. Like, mention Maria again, she's the proof. ▷ yields to S2

73.4–77.1 **S2** [turn] Yeah, good point. I keep forgetting that. Um, should I try again from the top? ▷ yields to S1

77.0–80.2 **S1** [turn] For sure. And honestly, you're sounding way more confident now. ▷ yields to S2

80.2–81.7 **S2** [turn] Thanks. That helps a lot. ▷ yields to S1

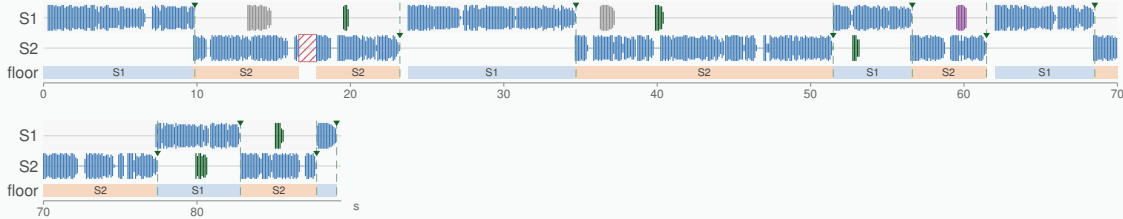
81.1–81.5 **S1** [backchannel] Yeah!

81.7–83.1 **S1** [turn] Ready when you are, Jay. ▷ yields to S2

83.0–86.9 **S2** [turn] Okay, here goes. Hi, my name is Jay, and I'm the co-founder of Fresco— ▷ yields, end

84.9–86.3 **S1** [laughter] Haha, alright, let's do it.

test split · tone: professional and business-like · noise: quiet apartment, 22.8 dB SNR · voices: S1 M, S2 M
 EOT 11 yields / 1 holds · INT 0 interruptions / 6 negatives (5 backchannels, 1 laughter) · 2 failed claims (excluded)



Transcript of test_00027

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.9 **S1** [turn] Alright, so for today's group activity, we're supposed to come up with three strategies for improving student participation. Did you, uh, have any ideas right off the bat? ▷ yields to S2

9.8–16.6 **S2** [turn] Actually, yes. I was thinking about using more cold-calling techniques, just to kind of, you know, bring in quieter students. But— ▷ holds, pauses

13.3–14.9 **S1** [failed claim] Right, like randomly calling names?

17.8–23.2 **S2** [turn] —yeah, exactly. Not in a way that puts anyone on the spot, but just to keep everyone mentally present. ▷ yields to S1

19.6–19.9 **S1** [backchannel] mm-hm

23.8–34.7 **S1** [turn] That's solid. I was also thinking, maybe some kind of points system? Like, low stakes, but you get a point every time you contribute, and then at the end there's some bonus, I don't know, extra credit or something. ▷ yields to S2

34.7–38.7 **S2** [turn] I do like that. But do you think students will just talk for the sake of points, or—

36.3–37.3 **S1** [failed claim] —yeah, like gaming it?

38.7–44.4 **S2** [turn] —yeah! Exactly. We'd have to make sure the contributions are actually meaningful, not just random comments.

39.9–40.5 **S1** [backchannel] totally

44.5–51.5 **S2** [turn] Maybe we could tie that in with peer feedback? Like, let students signal when they thought a peer's point actually moved the discussion forward. ▷ yields to S1

51.5–56.6 **S1** [turn] Oh, that's interesting. So it's more collaborative and less like, um, competition for airtime. ▷ yields to S2

52.8–53.2 **S2** [backchannel] right

56.5–61.5 **S2** [turn] And it could build in some self-regulation. They'll want to make good points, not just be loud. ▷ yields to S1

59.5–60.2 **S1** [laughter] haha true

62.0–68.5 **S1** [turn] Okay, so that's cold-calling, a points system with some peer feedback... We need one more strategy. You got something? ▷ yields to S2

68.5–77.4 **S2** [turn] What about using small group discussions before opening it up to the whole class? Sometimes that makes it easier for, uh, quieter folks to test their ideas. ▷ yields to S1

77.3–80.8 **S1** [turn] Yeah, like think-pair-share. That totally helps break the ice. I'm all for it.

79.9–80.7 **S2** [backchannel] exactly

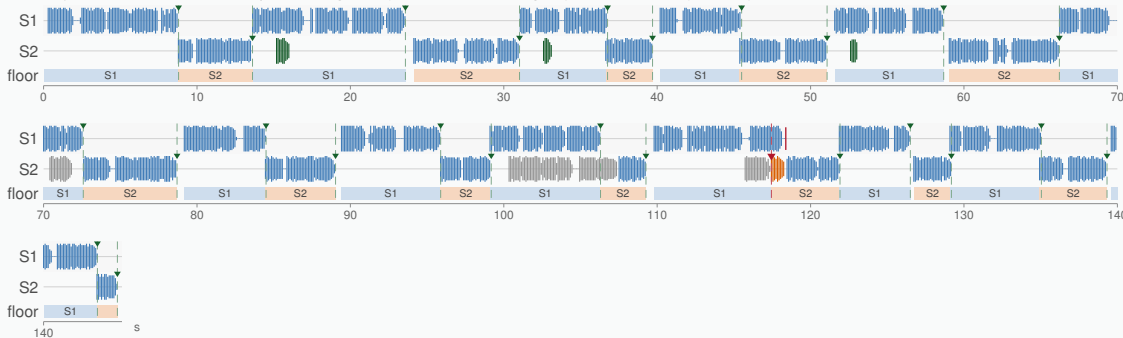
80.9–82.8 **S1** [turn] Alright, should we write these up for the board then? ▷ yields to S2

82.9–87.8 **S2** [turn] Yeah, let's do cold-calling, peer points with feedback, and then the think-pair-share thing. Sound good? ▷ yields to S1

85.1–85.7 **S1** [backchannel] perfect

87.8–89.1 **S1** [turn] Alright, I'll grab markers. ▷ yields, end

test split · tone: curious and inquisitive · noise: dorm room, 31.3 dB SNR; phone ring, door knock · voices: S1 M, S2 M
 EOT 25 yields / 0 holds · INT 1 interruptions / 3 negatives (3 backchannels, 0 laughter) · 3 failed claims (excluded)



Transcript of test_00035

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.8 **S1** [turn] Alright, so let's start mapping this out. We land in Tokyo at 11:20 a.m. on Tuesday, and check-in at the hotel isn't until, what, 3 p.m.? ▷ yields to S2

8.8–13.6 **S2** [turn] Yeah, uh, it's three. Should we, like, just leave our bags at the desk or try for early check-in, do you think? ▷ yields to S1

13.6–17.6 **S1** [turn] I think most places let you store your bags, so maybe we just do that. But—

15.2–16.0 **S2** [backchannel] Oh, right.

17.8–23.6 **S1** [turn] —do you want to go straight to Tsukiji for breakfast, or do something closer to the hotel first, since we'll have the luggage? ▷ yields to S2

24.1–31.0 **S2** [turn] Honestly, Tsukiji sounds fun, but with suitcases, maybe not. Maybe we just wander around Shinjuku? The park's supposed to be nice. ▷ yields to S1

31.1–36.8 **S1** [turn] Mm, good idea. Shinjuku Gyoen is right there. It's a pretty short walk, and maybe some coffee along the way. ▷ yields to S2

32.6–33.1 **S2** [backchannel] Right.

36.6–39.7 **S2** [turn] Should we book tickets for the park ahead or just get them at the gate? ▷ yields to S1

40.2–45.5 **S1** [turn] I think you can just show up. Unless it's cherry blossom season, but that's not for another couple of weeks. ▷ yields to S2

45.4–51.1 **S2** [turn] Okay, good. After that, maybe grab ramen for lunch? I saw a place called Ichiran on Google Maps nearby. ▷ yields to S1

51.6–54.4 **S1** [turn] Sounds perfect. I've been wanting to try Ichiran. Oh—

52.6–53.1 **S2** [backchannel] Yeah?

54.5–58.7 **S1** [turn] —what about that robot cafe thing? Is that nearby or the other direction? ▷ yields to S2

59.0–66.2 **S2** [turn] Um, that's actually just a few blocks from the hotel. But it gets booked up. Should I try to make a reservation for, like, later in the afternoon? $\triangleright$ yields to S1

66.2–72.6 **S1** [turn] Yeah, maybe around four, after we check in and drop off our stuff. Oh, and we'll probably crash pretty hard by dinner. $\triangleright$ yields to S2

70.4–71.9 **S2** [failed claim] Right, with the jet lag.

72.6–78.7 **S2** [turn] Should we just keep dinner super simple? Maybe grab convenience store food or go to that tonkatsu spot next door? $\triangleright$ yields to S1

79.2–84.5 **S1** [turn] Actually, convenience store food in Japan is kinda legendary. I'd be down for that, honestly. $\triangleright$ yields to S2

84.5–89.0 **S2** [turn] Same. Also, do you, uh, want to try hitting an onsen that night or wait till the next morning? $\triangleright$ yields to S1

89.4–95.9 **S1** [turn] I think I'd rather wait till morning. I'll probably just pass out early. But on Wednesday, let's definitely put that in. $\triangleright$ yields to S2

95.9–99.2 **S2** [turn] Wednesday morning onsen, then maybe hop the train to Hakone after? $\triangleright$ yields to S1

99.1–102.6 **S1** [turn] Yep, should we buy train tickets ahead, or just get them at the station?

100.3–107.4 **S2** [failed claim] Uh, let me check—I think with the Hakone Free Pass you can just activate it day-of. But maybe we want reserved seats for the Romancecar?

102.7–106.3 **S1** [turn] —oh, right, you're totally right. We should reserve that. Do you want me to look tonight? $\triangleright$ yields to S2

107.5–109.3 **S2** [turn] If you don't mind, that'd be great. $\triangleright$ yields to S1

109.8–118.4 **S1** [turn] No problem. Okay, so, Tuesday: land, hotel, park, ramen, robot cafe, konbini dinner, sleep. Wednesday: onsen, Hakone by train. That feel like the right flow? $\triangleright$ cut off

115.7–117.4 **S2** [failed claim] Yeah, yeah, I like it a lot.

117.4–118.3 **S2** [interruption] Oh, wait— $\triangleright$ takes floor from S1

118.5–121.9 **S2** [turn] —should we pack swimsuits for the onsen? Or do they give you stuff there? $\triangleright$ yields to S1

121.9–126.5 **S1** [turn] I think, uh, actually onsens are usually nude, so no swimsuits. They'll give us towels. $\triangleright$ yields to S2

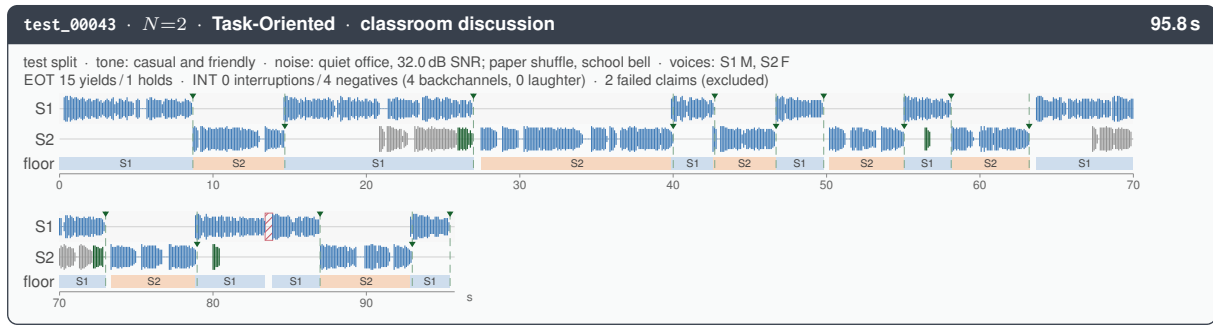
126.7–129.2 **S2** [turn] Ha, good thing you told me. I almost packed trunks. $\triangleright$ yields to S1

129.1–135.0 **S1** [turn] Haha, yeah, definitely easier to know that ahead. Anything else for Tuesday or Wednesday you want to squeeze in? $\triangleright$ yields to S2

134.9–139.3 **S2** [turn] No, I think that's a full start. We can always add stuff if we spot something cool. $\triangleright$ yields to S1

139.6–143.5 **S1** [turn] Alright, perfect. I'll jot that all down tonight and send you the itinerary. $\triangleright$ yields to S2

143.5–144.8 **S2** [turn] Awesome. Thanks, man. $\triangleright$ yields, end



Transcript of test_00043

start—end (s) speaker [ground-truth label] text $\triangleright$ outcome

0.3–8.7 **S1** [turn] Alright, so today we're supposed to come up with examples for the different types of chemical reactions, right? Like one synthesis, one decomposition, and so on. $\triangleright$ yields to S2

8.7–14.7 **S2** [turn] Yeah, and I think Ms. Patel wanted us to actually write out the balanced equations, not just list the types. $\triangleright$ yields to S1

14.6–23.5 **S1** [turn] Okay, cool, um, so for synthesis, maybe something super basic, like hydrogen gas plus oxygen gas gives water. H_2 plus O_2 gives H_2O .

20.9–25.9 **S2** [failed claim] Wait, you have to balance that, though. It's not just H_2 plus O_2 equals H_2O —

23.7–27.0 **S1** [turn] —oh right, it should be 2H_2 plus O_2 makes $2\text{H}_2\text{O}$. $\triangleright$ yields to S2

26.0–27.0 **S2** [backchannel] Yeah, exactly.

27.5–40.0 **S2** [turn] For decomposition, can we just use like, potassium chlorate breaking down into potassium chloride and oxygen? That's the one with the, um, KClO_3 becomes KCl plus O_2 . $\triangleright$ yields to S1

39.9–42.7 **S1** [turn] Yeah, that's classic. Wait, is that one balanced, though? $\triangleright$ yields to S2

42.6–46.7 **S2** [turn] No, you've gotta do 2KClO_3 makes 2KCl and 3O_2 . $\triangleright$ yields to S1

46.7–49.8 **S1** [turn] Right, right. I always forget the threes in front of O_2 . $\triangleright$ yields to S2

50.2–55.1 **S2** [turn] It trips everyone up. So, single replacement next? Iron and copper sulfate? $\triangleright$ yields to S1

55.1–58.1 **S1** [turn] Yeah! Like, Fe plus CuSO_4 makes FeSO_4 and Cu . $\triangleright$ yields to S2

56.4–56.8 **S2** [backchannel] Mm-hm.

58.2–63.2 **S2** [turn] I think that's already balanced? Fe plus CuSO_4 makes FeSO_4 plus copper, right. $\triangleright$ yields to S1

63.7–68.6 **S1** [turn] Yeah, that one's easy. Double replacement—um, let's do silver nitrate and sodium chloride?

67.3–72.1 **S2** [failed claim] Oh, but isn't it AgNO_3 plus NaCl gives AgCl and NaNO_3 ?

68.7–73.0 **S1** [turn] —yep, that's what I meant, sorry! Sometimes the formulas get me all twisted around. $\triangleright$ yields to S2

72.2–72.9 **S2** [backchannel] No worries.

73.4–79.0 **S2** [turn] Okay, last up is combustion. Are we just gonna go with methane? Like CH_4 plus O_2 ? $\triangleright$ yields to S1

78.9–83.4 **S1** [turn] Classic. So that'd be CH_4 plus 2O_2 gives CO_2 and $2\text{H}_2\text{O}$, right? $\triangleright$ holds, pauses

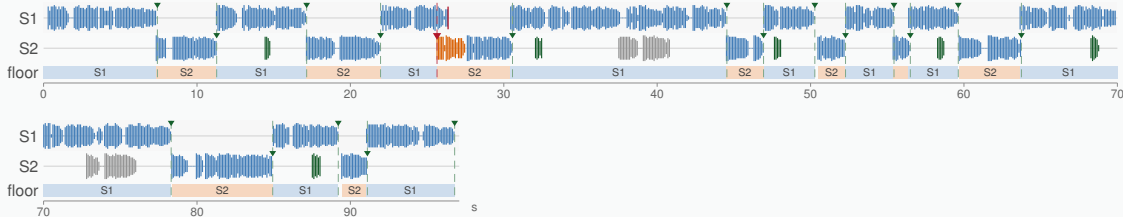
80.0–80.5 **S2** [backchannel] Right.

83.9–87.0 **S1** [turn] Haha, we nailed those. Should we list these out for the worksheet? $\triangleright$ yields to S2

87.0–93.0 **S2** [turn] Yeah, but let's double check the coefficients. I don't want to get marked down for like, missing a two somewhere. $\triangleright$ yields to S1

92.9–95.5 **S1** [turn] Good call. I'll write them out and you check as I go. $\triangleright$ yields, end

test split · tone: professional and business-like · noise: outdoor park, 35.1 dB SNR; chairs scraping, keyboard clatter · voices: S1 F, S2 F
EOT 18 yields / 0 holds · INT 1 interruptions / 6 negatives (6 backchannels, 0 laughter) · 2 failed claims (excluded)

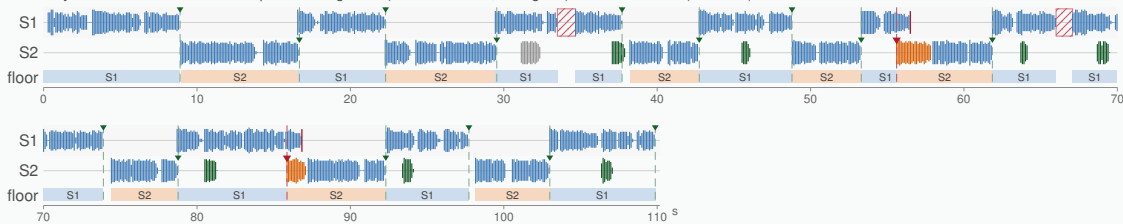


Transcript of test_00048

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–7.4 S1 [turn] Okay, let's get started. Our timeline for Saturday's park cleanup is tight, so I want to iron out roles and supply pickups first. ▷ yields to S2
7.3–11.3 S2 [turn] Makes sense. Do we know how many volunteers have confirmed for early setup? ▷ yields to S1
11.3–17.1 S1 [turn] Ten as of this morning. That's counting the two high school students, so we'll need an adult to supervise them. ▷ yields to S2
14.4–14.8 S2 [backchannel] Mm-hm.
17.1–22.0 S2 [turn] I can handle the student supervision, if you want to assign someone else to supply check-in? ▷ yields to S1
22.0–26.4 S1 [turn] Perfect. I'll put you as youth lead, then. For supply check-in, can we tap Michael? He's usually— ▷ cut off
25.6–27.5 S2 [interruption] He said he's unavailable before nine. ▷ takes floor from S1
27.6–30.6 S2 [turn] But I think Priya can cover that if she's still coming early? ▷ yields to S1
30.4–34.3 S1 [turn] Yeah, she RSVP'd for seven-thirty. I'll update that, thanks.
32.1–32.6 S2 [backchannel] Great.
34.4–40.4 S1 [turn] For tools, we're getting rakes and gloves from Public Works again. I'll pick those up Friday afternoon unless—
37.5–40.9 S2 [failed claim] Do you want backup on that? Last year you had to do two trips.
40.4–44.5 S1 [turn] —unless you prefer to split. Actually, that would help if your car is free. ▷ yields to S2
44.5–46.9 S2 [turn] I can meet you at the depot, then. Four-thirty? ▷ yields to S1
47.0–50.3 S1 [turn] Yeah, that works. I'll confirm the pickup slot and text you. ▷ yields to S2
47.6–48.1 S2 [backchannel] Thanks.
50.5–52.3 S2 [turn] Do we have coverage for the welcome table? ▷ yields to S1
52.3–55.4 S1 [turn] Sort of. I flagged that for Mark, but he hasn't emailed back. ▷ yields to S2
55.4–56.5 S2 [turn] Mm. Should I ping him again? ▷ yields to S1
56.4–59.6 S1 [turn] Please do. If we don't hear by Thursday, I'll shift someone else. ▷ yields to S2
58.3–58.7 S2 [backchannel] Sure.
59.7–63.7 S2 [turn] Are water and snacks set? I remember last time the delivery was late. ▷ yields to S1
63.7–69.9 S1 [turn] I called Whole Foods on Monday. They'll drop off at eight sharp. I'll get a backup case of water myself, just in case.
68.3–68.8 S2 [backchannel] Smart.
70.0–73.9 S1 [turn] Let's talk post-event. Do you want to run the sign-out table again? Or—
72.8–76.1 S2 [failed claim] I'm happy to, unless you think someone else should rotate in?
74.0–78.3 S1 [turn] —unless you're tired of it. But honestly you're best at the data piece, so that's ideal. ▷ yields to S2
78.4–84.9 S2 [turn] Happy to do it, then. Um, final thing, are we printing the new QR flyers or reusing last month's? ▷ yields to S1
85.0–89.2 S1 [turn] We'll print the new batch. I already uploaded the file to the drive, just use that link. ▷ yields to S2
87.5–88.1 S2 [backchannel] Perfect.
89.4–91.1 S2 [turn] I think that's everything on my list. ▷ yields to S1
91.1–96.8 S1 [turn] Same here. I'll send the updates by end of day so everyone's clear. Thanks for being on this—you keep us sane. ▷ yields, end

test split · tone: professional and business-like · noise: dorm room, 35.7 dB SNR · voices: S1 M, S2 M
EOT 15 yields / 2 holds · INT 2 interruptions / 7 negatives (7 backchannels, 0 laughter) · 1 failed claims (excluded)



Transcript of test_00091

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.9 S1 [turn] Alright, let's kick off with the group project outlines. The syllabus says everyone needs a clear hypothesis by Friday. So, what's your current proposal looking like, Mark? ▷ yields to S2
8.9–16.7 S2 [turn] Uh, yeah, so I was thinking about exploring how ambient noise impacts students' reading comprehension. I want to design a couple of short experiments for that. ▷ yields to S1
16.6–22.3 S1 [turn] That sounds promising. Are you planning to use surveys or more like direct observation in the classroom setting? ▷ yields to S2
22.3–29.5 S2 [turn] Actually, direct observation. I'd have, I guess, two groups—one with background noise and one in silence, then compare their test results. ▷ yields to S1
29.5–33.5 S1 [turn] Interesting. The variable control needs to be pretty tight, though. Otherwise— ▷ holds, pauses
31.1–32.4 S2 [failed claim] Yeah, that's my worry—
34.7–37.7 S1 [turn] —otherwise, you can't say which factor really made the difference. ▷ yields to S2
37.1–37.9 S2 [backchannel] Exactly.
38.2–42.7 S2 [turn] Um, what would you suggest for keeping it controlled? Like, same time of day for both groups? ▷ yields to S1
42.7–48.8 S1 [turn] Yeah, same time, identical materials, maybe even the same teacher if possible. Try to minimize variables. ▷ yields to S2
45.5–46.1 S2 [backchannel] Right.
48.8–53.3 S2 [turn] What about, like, the type of noise? Should it be just music, or like regular classroom chatter? ▷ yields to S1
53.3–56.5 S1 [turn] I'd say start simple: maybe white noise for one group, and then— ▷ cut off

55.6–57.9 **S2** [interruption] Or actual recorded classroom sounds? ▷ takes floor from S1

58.0–61.9 **S2** [turn] Because I feel like white noise is different from, uh... real disruptions, right? ▷ yields to S1

61.9–66.0 **S1** [turn] Oh, absolutely, yeah. Realistic noise would make your results more applicable. ▷ holds, pauses

63.7–64.2 **S2** [backchannel] mm-hm

67.1–73.9 **S1** [turn] But you might want to specify that in your limitations—like, the difference between controlled noise and unpredictable classroom stuff. ▷ yields to S2

68.7–69.5 **S2** [backchannel] Good point.

74.4–78.8 **S2** [turn] Are you expecting us to have preliminary data by next week, or just the design? ▷ yields to S1

78.7–83.0 **S1** [turn] Just the design for next week. Data collection comes after, once it's approved.

80.5–81.3 **S2** [backchannel] oh, okay

83.1–86.8 **S1** [turn] Did you have any concerns about the consent forms or the ethics review? ▷ cut off

85.9–87.1 **S2** [interruption] Um, yeah, actually— ▷ takes floor from S1

87.3–92.3 **S2** [turn] I was wondering if we need parental consent for kids under sixteen, or is school approval enough? ▷ yields to S1

92.3–97.7 **S1** [turn] You absolutely need parental consent, and the school has to sign off. I'll send you the template forms. ▷ yields to S2

93.4–94.1 **S2** [backchannel] Perfect.

98.2–103.0 **S2** [turn] Thanks for the feedback, that clears up a lot. Uh, anything else I should think about before Friday? ▷ yields to S1

103.0–109.9 **S1** [turn] That's about it. Just make sure your methodology section is detailed—sample size, timing, all of that. You're off to a solid start. ▷ yields, end

106.4–107.1 **S2** [backchannel] Thank you.

test_00093 · $N=2$ · Task-Oriented · board game night

63.0 s

test split · tone: enthusiastic and energetic · noise: school hallway, 35.0 dB SNR; notification chime · voices: S1 F, S2 F
EOT 16 yields / 1 holds · INT 1 interruptions / 4 negatives (2 backchannels, 2 laughter) · 3 failed claims (excluded)



Transcript of test_00093

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–5.6 **S1** [turn] Okay, so I set up most of the pieces, but I think I got the colors mixed up. Can you check if you're red or blue? ▷ yields to S2

5.6–8.6 **S2** [turn] I'm always red, come on, we've played this like twenty times. ▷ yields to S1

7.8–9.6 **S1** [laughter] Haha, right, force of habit.

9.7–15.1 **S1** [turn] Okay, so red for you. I'll—wait, which expansion did you want to use tonight? Because I see the dragons here too. ▷ yields to S2

15.2–17.9 **S2** [turn] Let's skip dragons, honestly, it slows the game down. ▷ yields to S1

16.7–18.3 **S1** [failed claim] Yeah, that's totally fair.

18.3–24.5 **S1** [turn] Alright, base game only, plus the trade routes expansion. Do you remember all the trade route rules, or should I do a quick refresher? ▷ yields to S2

24.9–26.9 **S2** [turn] Give me, like, the thirty second version? ▷ yields to S1

26.9–31.0 **S1** [turn] Uh, okay, so you can build a road through someone's settlement as long as you pay, um— ▷ holds, pauses

30.1–31.1 **S2** [failed claim] That's only if—

31.9–35.5 **S1** [turn] —as long as you pay two resources. And ports still work the same way. ▷ yields to S2

34.1–35.6 **S2** [failed claim] Got it, that's what I thought.

35.7–38.2 **S2** [turn] Alright, then let's just start with setup. Where's the dice? ▷ yields to S1

38.7–40.6 **S1** [turn] Oh, they rolled under the couch I think, hang on— ▷ cut off

39.9–42.1 **S2** [interruption] Hold up, I saw one by the bookshelf too! ▷ takes floor from S1

42.2–43.7 **S2** [turn] Ah, here it is. We're good.

43.2–44.3 **S1** [backchannel] Nice save!

43.8–47.3 **S2** [turn] Okay, so we'll put the settlements down first. Want to go first, or should I? ▷ yields to S1

47.2–50.0 **S1** [turn] Um, you go. I'm still sorting out the resource cards. ▷ yields to S2

49.9–52.2 **S2** [turn] Alright, starting on the wheat port as usual! ▷ yields to S1

52.6–55.1 **S1** [turn] Classic. I was hoping you'd switch it up just once. ▷ yields to S2

55.1–57.2 **S2** [turn] Why mess with a winning strategy, right? ▷ yields to S1

55.8–57.5 **S1** [laughter] Haha, I guess that's true.

57.6–60.0 **S1** [turn] Alright, I'll go brick and wood combo over here. ▷ yields to S2

58.3–58.9 **S2** [backchannel] Uh-huh.

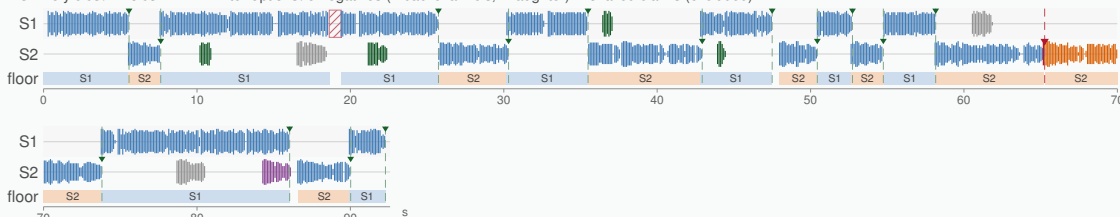
60.0–61.6 **S2** [turn] Nice, bold move. Are we set? ▷ yields to S1

61.5–62.7 **S1** [turn] Yep, let's do this! ▷ yields, end

test_00157 · $N=2$ · Task-Oriented · classroom discussion

92.6 s

test split · tone: casual and friendly · noise: quiet office, 29.3 dB SNR; notification chime; distant speech · voices: S1 F, S2 M
EOT 15 yields / 1 holds · INT 1 interruptions / 5 negatives (4 backchannels, 1 laughter) · 3 failed claims (excluded)



Transcript of test_00157

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–5.6 **S1** [turn] Okay, let's just start with number four on the worksheet because that one seemed to throw everyone off yesterday. ▷ yields to S2

5.5–7.6 **S2** [turn] Oh yeah, that was the graph question, right? ▷ yields to S1

7.6–12.0 **S1** [turn] Exactly. The one with population growth over time, and it asked for carrying capacity.

10.2–11.0 **S2** [backchannel] Right, right.

12.1–18.6 **S1** [turn] So, the main thing is you have to look for where the curve levels out. That's your carrying capacity, it's not where it starts or where it spikes up— **> holds, pauses**

16.5–18.4 **S2** [failed claim] But it like, dips down a bit at the end—

19.4–25.7 **S1** [turn] —yeah, that's normal. Sometimes populations fluctuate around the carrying capacity, it's not always a perfectly flat line. **> yields to S2**

21.2–22.4 **S2** [backchannel] Oh! Okay, cool.

25.8–30.3 **S2** [turn] So, we'd just say it's whatever value it's more or less hovering at near the right edge of the graph? **> yields to S1**

30.2–35.5 **S1** [turn] Yep. In this case I think it was, like, eight hundred? I'm not looking at it right now, but I remember it hovering there. **> yields to S2**

35.5–37.6 **S2** [turn] Yeah! Eight hundred, that's what I wrote too.

36.5–37.2 **S1** [backchannel] Perfect.

37.7–42.9 **S2** [turn] Haha, I actually almost circled nine hundred at first, but then I noticed it kept dropping back down. **> yields to S1**

42.8–47.5 **S1** [turn] Yeah, that's what tricks people—just because it goes a little higher doesn't mean that's the steady state. **> yields to S2**

43.9–44.5 **S2** [backchannel] Right.

48.0–50.4 **S2** [turn] Okay, so do you want to tackle number five now? **> yields to S1**

50.4–52.7 **S1** [turn] That's the one about limiting factors, right? **> yields to S2**

52.6–54.7 **S2** [turn] Uh, yeah, with the rabbits and the foxes. **> yields to S1**

54.8–58.2 **S1** [turn] Okay, so what'd you put for the main limiting factor in that scenario? **> yields to S2**

58.1–65.2 **S2** [turn] I wrote predation, 'cause the fox population was increasing at the same time the rabbits started decreasing, like right in that table. **> cut off**

60.5–61.9 **S1** [failed claim] Yup. That makes sense.

65.3–70.0 **S2** [interruption] But is that what Mrs. Kim was actually looking for? I remember she said something about drought... **> takes floor from S1**

70.0–73.8 **S2** [turn] I mean, I didn't see evidence for that in the numbers, but the homework mentioned it. **> yields to S1**

73.8–80.2 **S1** [turn] No, you're right. If the question didn't say anything about a drought in the data they gave us, then predation is the most direct answer.

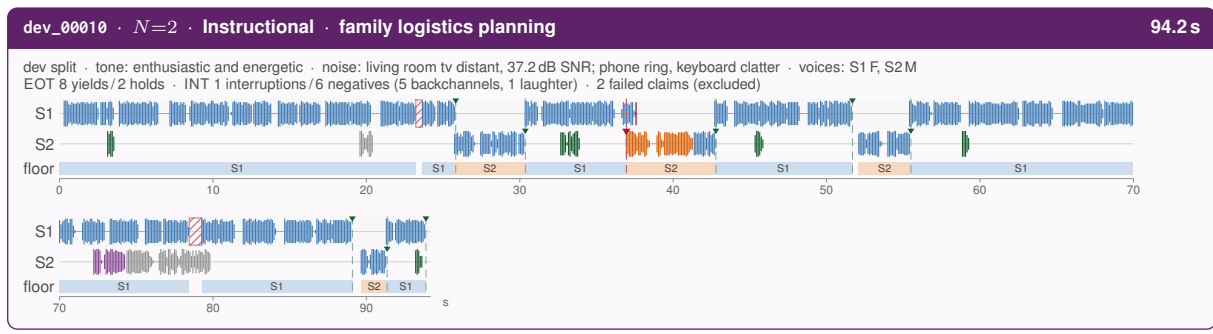
78.7–80.6 **S2** [failed claim] Okay, that makes me feel better.

80.3–86.0 **S1** [turn] Honestly, they're just checking if we're reading what's actually there, not just guessing based on something random she said last week. **> yields to S2**

84.3–86.1 **S2** [laughter] Haha, yeah, that's totally true.

86.6–90.0 **S2** [turn] Alright, you wanna move on to number six, or do you feel good about those? **> yields to S1**

89.9–92.3 **S1** [turn] Let's hit six, just in case. It never hurts. **> yields, end**



Transcript of dev_00010

start—end (s) speaker [ground-truth label] text **> outcome**

0.3–12.6 **S1** [turn] Alright, so if we want Sunday to go smoothly, we need a plan—I swear it'll only take a few minutes to run through. I'll set up a schedule, and I'll show you how we can keep track of who's doing what, without total chaos.

3.1–3.6 **S2** [backchannel] Uh-huh.

12.7–23.2 **S1** [turn] So first—breakfast. I'll handle pancakes, you're on juice and coffee, and then afterwards we'll divide and conquer on packing the car. I'll write all this on the whiteboard so it's visible. **> holds, pauses**

19.6–20.5 **S2** [failed claim] Wait, so—

23.7–25.8 **S1** [turn] —you want to switch? You don't want to do coffee? **> yields to S2**

25.8–30.4 **S2** [turn] I was just gonna ask, do I need to make both or can it be just set out, y'know? **> yields to S1**

30.4–37.6 **S1** [turn] Oh, good question! No, the coffee machine's automatic, so just double-check the water and punch the button at seven sharp. For juice, just pour it. Easy. **> cut off**

32.7–33.9 **S2** [backchannel] Cool, that's—yeah.

36.9–41.3 **S2** [interruption] And what about the kids' stuff? Like, am I supposed to check their bags? **> takes floor from S1**

41.4–42.8 **S2** [turn] Or is that on you? **> yields to S1**

42.7–51.7 **S1** [turn] Great you mentioned that! I'm making a checklist for each kid—clothes, toys, sunscreen, you know the drill. You just double-check Mark's backpack and I've got Lily's, okay? **> yields to S2**

45.3–45.9 **S2** [backchannel] Right.

52.1–55.5 **S2** [turn] So, like, when do we load up? After breakfast or later? **> yields to S1**

55.4–68.1 **S1** [turn] Yes! Right after eating—otherwise stuff always gets missed. We'll do a final sweep of the living room for stray shoes and Lego, and then load everything right into the trunk. I'll put the blanket in last so it's easy to grab.

58.9–59.3 **S2** [backchannel] mm-hm.

68.1–78.5 **S1** [turn] And then, before we leave, you do a seatbelt and snack check. I'll handle bathroom runs. That way no meltdown in the driveway. Oh, and playlist! I'll cue up the playlist for the car. **> holds, pauses**

72.2–74.3 **S2** [laughter] Haha, that's actually important.

74.4–79.9 **S2** [failed claim] So I just have to make sure Mark's not got, like, rocks in his pockets and he's buckled?

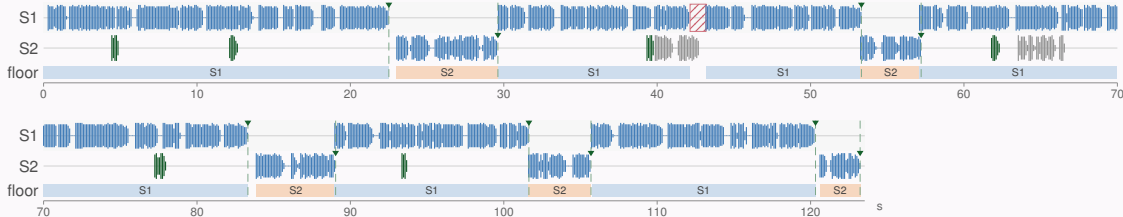
79.3–89.1 **S1** [turn] —Exactly, no rock collection this time. That's the last step and then we pull out at eight. If we stick to this, we'll be on the road—fingers crossed—before the whole neighborhood does. **> yields to S2**

89.7–91.4 **S2** [turn] Alright, sounds like a plan. **> yields to S1**

91.3–93.9 **S1** [turn] Awesome! Seriously, we'll be like a well-oiled machine. **> yields, end**

93.2–93.7 **S2** [backchannel] Yep.

dev split · tone: casual and friendly · noise: quiet office, 26.1 dB SNR; paper shuffle · voices: S1 M, S2 M
EOT 10 yields / 1 holds · INT 0 interruptions / 6 negatives (6 backchannels, 0 laughter) · 2 failed claims (excluded)

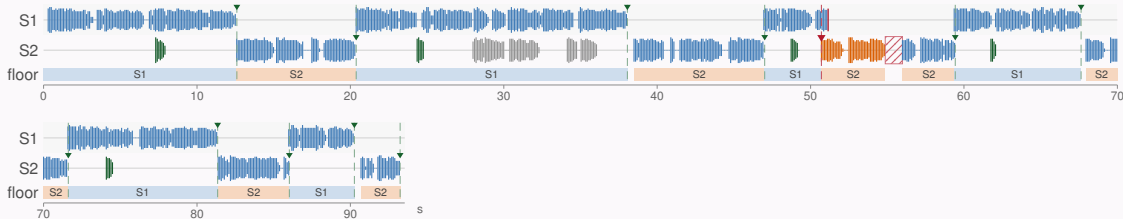


Transcript of dev_00047

start-end (s) speaker [ground-truth label] text ▷ outcome

- 0.3–8.9 S1 [turn] Alright, so today I thought we could start by breaking down how inflation actually works, since it's basically been in every headline this week.
- 4.5–4.9 S2 [backchannel] Uh-huh.
- 8.9–22.5 S1 [turn] So, in the simplest terms, inflation just means that the average price of stuff—goods and services, basically—goes up over time. But it gets weird because not everything gets more expensive at the same rate. ▷ yields to S2
- 12.1–12.7 S2 [backchannel] Right.
- 23.0–29.6 S2 [turn] So is that, like, when I go buy coffee and suddenly it costs fifty cents more than last month, that's inflation? ▷ yields to S1
- 29.7–42.1 S1 [turn] Exactly. That's literally it—you're seeing the effect right there. But, um, more officially, economists look at these giant baskets of goods—hundreds of items—and track how the prices change as a whole. ▷ holds, pauses
- 39.3–39.8 S2 [backchannel] Uh-huh.
- 39.9–42.7 S2 [failed claim] Hang on, basket? Like an actual basket—
- 43.2–53.3 S1 [turn] —not a literal basket, no. It's a metaphor. They just mean a collection of lots of typical things that people buy—everything from bread to rent, even movie tickets. ▷ yields to S2
- 53.3–57.2 S2 [turn] Oh, okay. That would've been, like, a wild grocery trip. ▷ yields to S1
- 57.1–67.8 S1 [turn] Haha, yeah, exactly. So when you hear something like "inflation rose three percent last year," it means the overall price of that whole basket went up by three percent.
- 61.8–62.3 S2 [backchannel] Right.
- 63.5–66.6 S2 [failed claim] But why, uh, is it bad if it goes up? Or—
- 67.9–83.3 S1 [turn] —it kind of depends. Some inflation is normal, even good, because it means the economy's moving and people are buying things. But when it jumps too fast, wages can't keep up and it makes everything feel more expensive, especially essentials. ▷ yields to S2
- 77.2–78.0 S2 [backchannel] Makes sense.
- 83.9–89.0 S2 [turn] So how do they even measure that? Like, isn't everyone's shopping basket different? ▷ yields to S1
- 89.0–101.6 S1 [turn] Good question. They basically average it out. They survey what people buy across the country, weight common stuff like rent and food more heavily, and use all that to build, well, a sort of 'average' basket. ▷ yields to S2
- 93.4–93.8 S2 [backchannel] Mmm.
- 101.6–105.7 S2 [turn] And is that what the, uh, CPI is? I see that all the time. ▷ yields to S1
- 105.7–120.3 S1 [turn] Yeah, that's exactly it. CPI stands for Consumer Price Index. It's the main number reporters talk about when they mention inflation. It's not perfect, but it gives us a decent picture of what's happening with prices overall. ▷ yields to S2
- 120.6–123.2 S2 [turn] Cool. Thanks for the, uh, crash course. ▷ yields, end

test split · tone: curious and inquisitive · noise: quiet apartment, 30.8 dB SNR; paper shuffle, keyboard clatter; distant speech · voices: S1 F, S2 F
EOT 11 yields / 1 holds · INT 1 interruptions / 5 negatives (5 backchannels, 0 laughter) · 2 failed claims (excluded)



Transcript of test_00015

start-end (s) speaker [ground-truth label] text ▷ outcome

- 0.3–12.6 S1 [turn] Alright, so if you all want to figure out where to eat with a big group, the most efficient way is to set a few ground rules first. Like, you need to find out if anyone has dietary restrictions or allergies before you even start picking places. ▷ yields to S2
- 7.3–7.9 S2 [backchannel] Oh, sure.
- 12.6–20.4 S2 [turn] But what if someone forgets to mention an allergy until you're already at the restaurant? I mean, people don't always remember right away. ▷ yields to S1
- 20.4–30.2 S1 [turn] That's true, but that's why you double-check in the group chat before finalizing the choice. Just ask, 'Hey, anyone need gluten-free or vegan options?'—get it out in the open. It saves time.
- 24.4–24.8 S2 [backchannel] Right.
- 28.0–32.3 S2 [failed claim] So after you get everyone's restrictions, do you just pick one place or do you—
- 30.3–38.1 S1 [turn] —you give two or three options, tops. Any more than that and nobody can decide. You have to keep it limited or the conversation drags on forever. ▷ yields to S2
- 34.1–36.1 S2 [failed claim] Yeah, that's... that's totally true.
- 38.5–47.0 S2 [turn] But what if the group splits on the options? Like, three people want sushi and three want Mexican. How do you keep that from turning into a whole debate? ▷ yields to S1
- 47.0–51.1 S1 [turn] If it's split, you do a quick poll—don't let it turn into an argument. And if it's still even, you just flip a coin. Someone has to make a call or you'll never eat. ▷ cut off
- 48.7–49.2 S2 [backchannel] Okay.
- 50.7–54.9 S2 [interruption] But say you have a group of ten, and the restaurant can't seat that many right away— ▷ takes floor from S1 ▷ holds, pauses
- 56.0–59.4 S2 [turn] —do you call ahead, or is it better to just show up and hope for the best? ▷ yields to S1
- 59.4–67.6 S1 [turn] Always call ahead or book online if you can. Most places want heads-up for big groups. Shows respect for their staff, and you're less likely to wait

around. ▷ yields to S2

61.8–62.2 **S2** [backchannel] Mm-hm.

68.0–71.6 **S2** [turn] What about splitting the bill? That always gets awkward in big groups. ▷ yields to S1

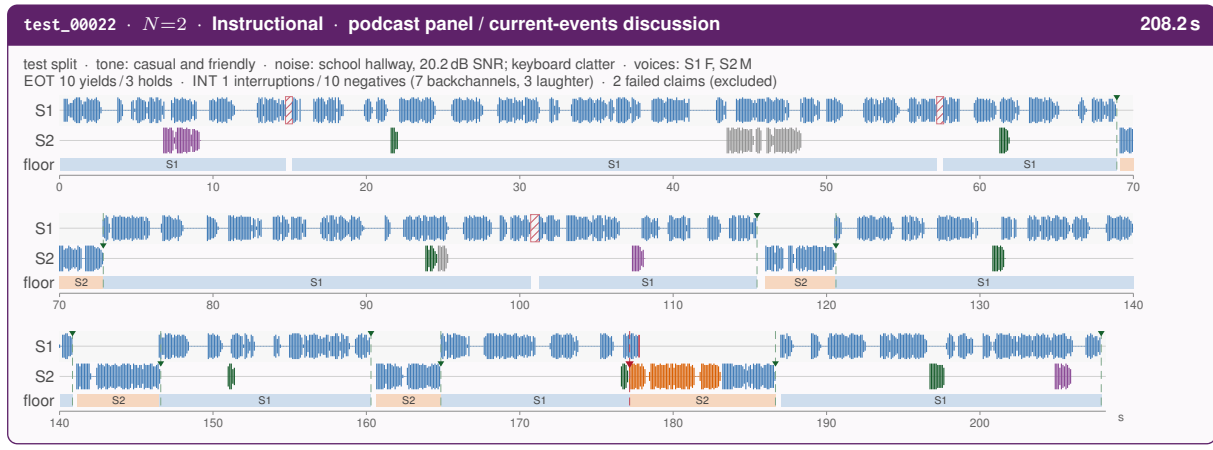
71.6–81.3 **S1** [turn] Yeah, it does. Best way is to tell your server at the start if people want separate checks. Otherwise, use an app like Splitwise or just have one person pay and everyone Venmos after. ▷ yields to S2

74.1–74.6 **S2** [backchannel] Uh-huh.

81.4–86.0 **S2** [turn] So basically the main thing is just plan as much as you can before you meet up, right? ▷ yields to S1

86.0–90.3 **S1** [turn] Exactly. The less you leave to the last minute, the smoother it goes for everyone. ▷ yields to S2

90.7–93.2 **S2** [turn] Great, thanks. That actually clears up a lot. ▷ yields, end



Transcript of test_00022

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–14.7 **S1** [turn] Alright, so today we're diving into the, uh, surprisingly thrilling topic of digital privacy laws and how they're changing right now. Don't all cheer at once. ▷ holds, pauses

6.8–9.2 **S2** [laughter] Haha, yeah, I've been waiting all week for this one.

15.2–35.7 **S1** [turn] You joke, but it's actually kind of a big deal. So, right now, governments around the world are updating, or sometimes just inventing, laws around things like data collection, cookies, and what companies can do with your info.

21.6–22.1 **S2** [backchannel] Right

35.8–57.2 **S1** [turn] So, the big headline recently was the EU's Digital Services Act. Think of it as GDPR's more ambitious cousin—lots of new requirements on transparency and what platforms have to take down when it comes to misinformation. ▷ holds, pauses

43.5–48.4 **S2** [failed claim] Wait, so when you say platforms, like, you mean, uh, basically anything with a comment section?

57.6–68.9 **S1** [turn] —yeah, pretty much. Social media sites, some news websites, even online marketplaces. So, Etsy sellers, beware. ▷ yields to S2

61.3–61.9 **S2** [backchannel] Oh wow

69.1–72.8 **S2** [turn] So, like, what's something I might see change, as just, you know, a random user? ▷ yields to S1

72.8–100.7 **S1** [turn] Okay, so you'll probably see way more pop-ups, which is honestly thrilling. But also, more actual explanations—like, instead of just 'we use cookies,' you might get told exactly what the cookies do. Or, if you post something flagged as misinformation, you might get a specific notice rather than a generic warning. ▷ holds, pauses

93.9–94.6 **S2** [backchannel] Makes sense

94.7–95.3 **S2** [failed claim] But, uh—

101.3–115.5 **S1** [turn] —oh, and companies have to make it way easier to report illegal stuff. There's now, like, forms for complaints, which I'm sure will never ever be abused. ▷ yields to S2

107.3–108.2 **S2** [laughter] Haha, right

116.0–120.6 **S2** [turn] So does this even apply to, uh, U.S.-based companies when I'm in, like, California? ▷ yields to S1

120.5–140.8 **S1** [turn] Sort of. If a company operates in the EU or has EU users, then they have to follow the rules for those folks. But in the States, California does have its own law, CCPA, which is like GDPR on a budget. ▷ yields to S2

130.8–131.7 **S2** [backchannel] Ah, okay

141.1–146.6 **S2** [turn] Is it true that, like, companies can get sued if they screw up the notifications or data collection stuff? ▷ yields to S1

146.5–160.3 **S1** [turn] Yeah, and the fines are not small. In the EU, it's literally billions. Like, the kind of change that could buy you a fleet of private jets—if you're into that. ▷ yields to S2

151.0–151.5 **S2** [backchannel] Yikes

160.7–164.9 **S2** [turn] Are there any, uh, exceptions? Like, if you're a tiny business, do you still get hit? ▷ yields to S1

164.9–177.8 **S1** [turn] Yeah, there are thresholds—like, if you have under a certain number of employees or users, you might get less strict rules. But honestly, most policies were written assuming Mark Zuckerberg is listening, not, like, your local coffee shop. ▷ cut off

176.6–177.1 **S2** [backchannel] Yeah

177.2–183.1 **S2** [interruption] Okay, last thing—do you actually think this stuff helps, like, the average person? Or is it just more paperwork? ▷ takes floor from S1

183.2–186.7 **S2** [turn] Because I feel like people just click through cookie banners no matter what. ▷ yields to S1

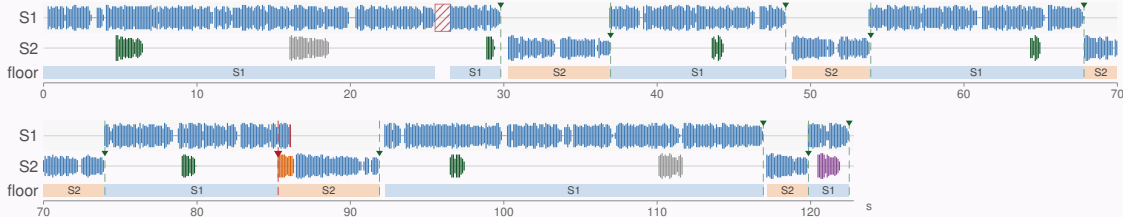
187.0–202.4 **S1** [turn] You're not wrong. I mean, some transparency is better than none, but most folks just want the pop-ups to go away. So, it's progress, but, like, frustrating progress.

196.7–197.7 **S2** [backchannel] Yeah, absolutely

202.5–207.9 **S1** [turn] If the next law could just outlaw all cookie banners, I'd vote for it. ▷ yields, end

204.9–206.0 **S2** [laughter] Haha, please, yes

test split · tone: professional and business-like · noise: office hum, 20.9 dB SNR; paper shuffle, school bell · voices: S1 M, S2 M
EOT 10 yields / 1 holds · INT 1 interruptions / 7 negatives (6 backchannels, 1 laughter) · 2 failed claims (excluded)



Transcript of test_00059

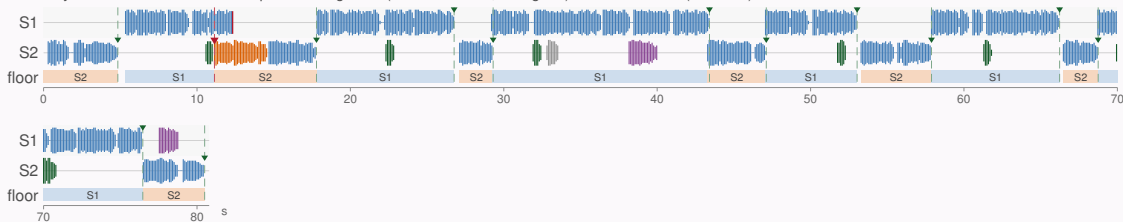
start-end (s) speaker [ground-truth label] text > outcome

- 0.3–11.0 **S1** [turn] Okay, so uh, thanks for having me on today. We—um, we’re going to look at how the latest economic data actually reflects on, you know, the wider global market, if that’s all right?
- 4.7–6.5 **S2** [backchannel] Absolutely, sounds good.
- 11.2–25.5 **S1** [turn] Um, so first, the U.S. jobs report came out last Friday and showed, like, strong hiring but also that wage growth is slowing down a little. Which uh, economists say is actually kind of reassuring for inflation, given, you know— > holds, pauses
- 16.1–18.6 **S2** [failed claim] Right, so slowing wage growth can be a good thing—
- 26.5–29.8 **S1** [turn] —that we’ve been dealing with rising prices the last couple years. > yields to S2
- 28.9–29.4 **S2** [backchannel] Got it.
- 30.3–37.0 **S2** [turn] So does that mean the Fed is, uh, more likely to cut rates soon? I mean, a lot of people seem to be hoping for that. Just to, you know— > yields to S1
- 36.9–48.4 **S1** [turn] Yeah, so um, not immediately. The Fed wants to see like several months in a row, uh, confirming that inflation is slowing steadily before they change rates. They’re really cautious right now. > yields to S2
- 43.6–44.3 **S2** [backchannel] Makes sense.
- 48.8–53.9 **S2** [turn] So, um, looking at the global side, how’s Europe doing? I heard Germany was having some problems— > yields to S1
- 53.8–67.8 **S1** [turn] Yeah, Germany’s economy actually shrank a bit this quarter, mostly due to weak exports and, uh, energy costs. But the European Central Bank hasn’t really reacted with big changes yet. They’re also sort of in a wait-and-see mode. > yields to S2
- 64.3–65.0 **S2** [backchannel] Okay.
- 67.9–74.0 **S2** [turn] And, uh, Asia? China especially—I keep seeing headlines about their property market. Is it as bad as everyone says? > yields to S1
- 74.0–86.1 **S1** [turn] Um, yeah, China’s real estate sector is still basically in crisis. Several big developers have defaulted, and confidence is pretty low. The government’s tried some stimulus, but um, so far it’s only, it’s only stabilized things a little. > cut off
- 79.0–80.0 **S2** [backchannel] Ah, okay.
- 85.3–86.4 **S2** [interruption] So, can I just ask— > takes floor from S1
- 86.5–91.9 **S2** [turn] What should, uh, the average listener actually watch for in terms of market trends over the next, say, three months? > yields to S1
- 92.2–103.9 **S1** [turn] That’s a great question. So, I’d say watch the inflation numbers and anything about central bank rate changes, especially from the Fed. If inflation keeps slowing, markets could respond with rallies.
- 96.5–97.5 **S2** [backchannel] Right, rallies.
- 104.0–116.9 **S1** [turn] And just, um, be cautious with news about China. If the property market worsens again, it could ripple out into other economies. So, uh, global headlines are important to follow—even if you’re just investing here at home. > yields to S2
- 110.1–111.7 **S2** [failed claim] Interesting. Yeah, good advice.
- 117.1–119.9 **S2** [turn] Thanks for breaking that down, by the way. Super clear. > yields to S1
- 119.9–122.5 **S1** [turn] Oh, uh, yeah, my pleasure. Happy to help where I can. > yields, end
- 120.5–121.9 **S2** [laughter] Haha, thanks again.

test_00076 · N=2 · Instructional · roommate household planning

80.8 s

test split · tone: enthusiastic and energetic · noise: school hallway, 34.2 dB SNR · voices: S1 M, S2 F
EOT 12 yields / 0 holds · INT 1 interruptions / 8 negatives (6 backchannels, 2 laughter) · 1 failed claim (excluded)



Transcript of test_00076

start-end (s) speaker [ground-truth label] text > outcome

- 0.3–4.8 **S2** [turn] Okay, let’s get started. Can you walk me through your glorious new house chore system? > yields to S1
- 5.3–12.3 **S1** [turn] Uh, yeah, sure! So, um, basically I thought we could have, uh, a rotating schedule—like each week we swap chores, so no one gets stuck with the bathroom twice in a row, you know? > cut off
- 10.6–11.1 **S2** [backchannel] mm-hm
- 11.1–14.6 **S2** [interruption] Wait, do you mean literally every chore changes hands every week? > takes floor from S1
- 14.7–17.8 **S2** [turn] Or just the big ones like kitchen and bathroom get swapped? > yields to S1
- 17.8–26.8 **S1** [turn] Oh, yeah, sorry! Uh, just the main ones—so kitchen, bathroom, taking out trash. Stuff like watering plants or, um, organizing the mail would be, like, add-ons if you want. > yields to S2
- 22.3–22.9 **S2** [backchannel] right
- 27.1–29.3 **S2** [turn] And how do we know who’s got what each week? > yields to S1
- 29.2–36.4 **S1** [turn] So, um, I was thinking a whiteboard by the fridge? I can write everyone’s name under their chores on Sundays, like after grocery shopping.
- 31.9–32.5 **S2** [backchannel] right
- 32.8–33.6 **S2** [failed claim] Hang on—
- 36.5–43.4 **S1** [turn] —I can make a template so it’s fast. I’ll even draw little doodles next to each chore. I know that’s, like, dorky, but it helps me remember. > yields to S2
- 38.2–40.1 **S2** [laughter] Haha, I actually love that.
- 43.3–47.1 **S2** [turn] So, if I forget to check the board and totally space out, what do I do? > yields to S1

47.1–53.0 **S1** [turn] Oh, you can just text me or, like, ask in the group chat. I'll update reminders there too, but the board is official. ▷ yields to S2

51.8–52.4 **S2** [backchannel] uh-huh

53.3–57.9 **S2** [turn] And what if someone forgets to do their chore? Like, the bathroom doesn't get cleaned one week? ▷ yields to S1

57.9–66.2 **S1** [turn] Um... so I think we just remind each other—nicely? No shaming! And, uh, if it keeps happening, we can figure out a consequence, but I feel like we're all responsible. ▷ yields to S2

61.3–61.8 **S2** [backchannel] yeah

66.5–68.8 **S2** [turn] Do you want to do a test run this week then? ▷ yields to S1

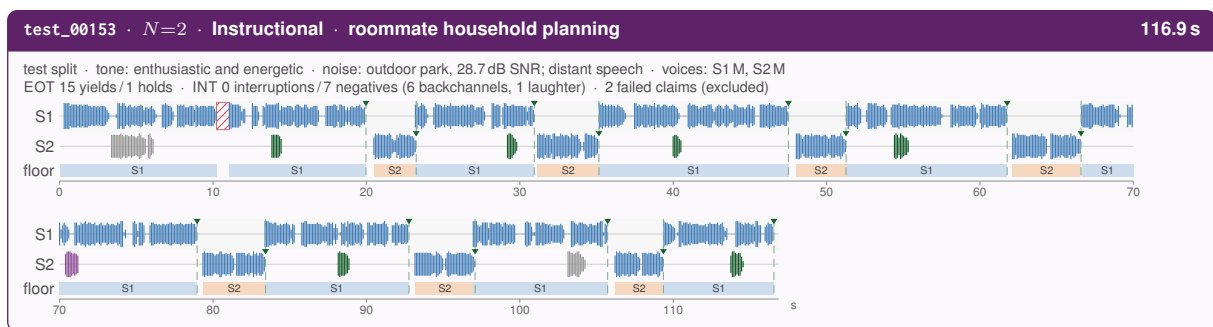
68.7–72.2 **S1** [turn] Yes! Uh, can we start tonight? I already have a marker for the board.

69.9–70.9 **S2** [backchannel] Perfect.

72.3–76.5 **S1** [turn] Awesome. Thank you for, uh, letting me over-explain. I know I get really into this stuff. ▷ yields to S2

76.5–80.5 **S2** [turn] Seriously, it's way better than chaos. Thanks for organizing us! ▷ yields, end

77.5–78.8 **S1** [laughter] Haha, anytime.



Transcript of test_00153

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–10.2 **S1** [turn] Alright, so let's get started on the roommate schedule! I promise this will make our life way easier, and honestly, it can even be kind of fun once you get into it. ▷ holds, pauses

3.4–6.2 **S2** [failed claim] Uh, yeah, okay, but like, do you mean just chores? Or—

11.1–20.0 **S1** [turn] —not just chores! I mean, we can map out stuff like cleaning, grocery runs, even, like, who's in charge of dinner if we feel ambitious. ▷ yields to S2

13.8–14.5 **S2** [backchannel] Oh wow

20.5–23.2 **S2** [turn] So, uh, how do we start? Like, do we write it out or— ▷ yields to S1

23.2–30.9 **S1** [turn] Great question! I actually think a whiteboard on the fridge is perfect for this. That way, no one can say they didn't see it, you know? ▷ yields to S2

29.2–29.8 **S2** [backchannel] Right

31.2–35.2 **S2** [turn] But do we have to, like, stick to the same chores every week? Or does it rotate? ▷ yields to S1

35.2–47.5 **S1** [turn] I actually recommend rotating! That way, nobody gets stuck always scrubbing the bathroom. For example, we could assign, say, dishes, vacuuming, and trash, then just swap every Sunday night. ▷ yields to S2

40.0–40.6 **S2** [backchannel] mm-hm

48.0–51.3 **S2** [turn] But, uh, what if I can't do mine on the day it's supposed to be done? ▷ yields to S1

51.3–61.8 **S1** [turn] Totally normal! Just let everyone know—you can swap a day, or trade with someone if they're cool with it. The key thing is communication, so nobody gets blindsided. ▷ yields to S2

54.4–55.4 **S2** [backchannel] Makes sense

62.1–66.6 **S2** [turn] Wait, so, what if, like, I'm really bad at cooking? Do I still have to do dinner sometimes? ▷ yields to S1

66.6–79.0 **S1** [turn] You know, honestly, we can play to our strengths! If you hate cooking, maybe you can double up on something else and someone else can swap. Or, hey, frozen pizza is always a valid dinner option. ▷ yields to S2

70.4–71.3 **S2** [laughter] haha, phew

79.3–83.4 **S2** [turn] How do we remember whose turn it is, though? Like, I always forget stuff like that. ▷ yields to S1

83.4–92.8 **S1** [turn] That's where the whiteboard is genius! We'll write names right on it for each task, and then every Sunday night, after dinner or whatever, we wipe and rotate. ▷ yields to S2

88.2–88.9 **S2** [backchannel] Oh, cool

93.2–97.1 **S2** [turn] Uh, do we need a reminder or something? I mean, what if we forget to rotate? ▷ yields to S1

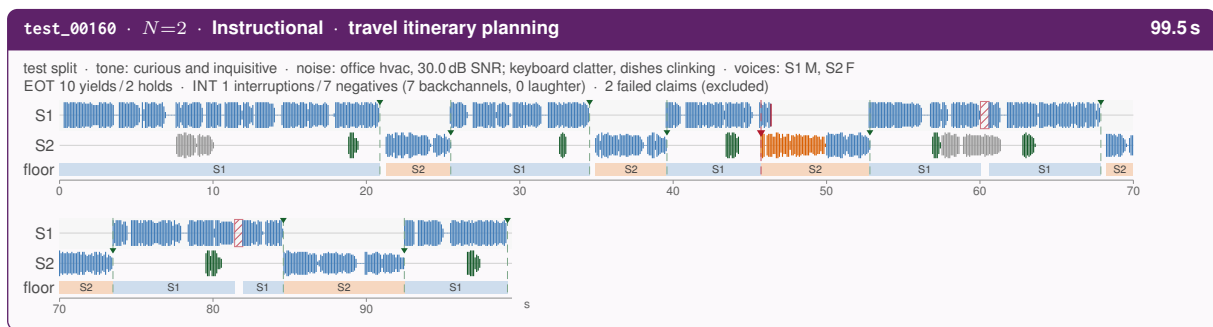
97.0–105.7 **S1** [turn] Ha! We could set a phone alarm for Sunday night. Or, like, make it part of our weekly movie night—just do the switch before the movie starts. ▷ yields to S2

103.1–104.3 **S2** [failed claim] Yeah, that would help

106.2–109.4 **S2** [turn] So, um, can we try it this week? I mean, just like, as a test? ▷ yields to S1

109.4–116.6 **S1** [turn] Absolutely! I'll grab a marker today, and we can fill in the first round. Honestly, I think we'll love it—you'll see. ▷ yields, end

113.8–114.6 **S2** [backchannel] Sounds good



Transcript of test_00160

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–12.7 **S1** [turn] Alright, so the basic idea for the trip is to start in Barcelona, then hop over to Valencia, and finally end up in Madrid. I think three nights in Barcelona, two in Valencia, and three in Madrid is pretty balanced.

7.6–10.0 **S2** [failed claim] Three in Barcelona feels like a lot—

12.8–20.9 **S1** [turn] —yeah, but there's a lot there. Sagrada Familia, the Gothic Quarter, beaches, you know. And it gives us a little buffer in case of jet lag the first day. > yields to S2

18.9–19.5 **S2** [backchannel] Right

21.3–25.5 **S2** [turn] So what's the plan for getting from Barcelona to Valencia? Train or bus? > yields to S1

25.5–34.6 **S1** [turn] Train, definitely. The high-speed AVE takes about three hours, and it's much more comfortable than the bus. I can book us tickets in advance for a morning departure. > yields to S2

32.6–33.1 **S2** [backchannel] mm-hm

34.9–39.6 **S2** [turn] Can we fit in the Albufera lake thing while we're in Valencia? Or is that too much? > yields to S1

39.5–46.4 **S1** [turn] No, that's doable. We can do the City of Arts and Sciences one morning, then Albufera in the afternoon. It's a short bus ride out, maybe forty minutes. > cut off

43.5–44.4 **S2** [backchannel] Oh, great

45.7–49.9 **S2** [interruption] Wait, do you know if the paella places out there are open in the evening or just lunch? > takes floor from S1

50.0–52.8 **S2** [turn] Because if it's only lunch, maybe we swap the order. > yields to S1

52.8–60.0 **S1** [turn] Good point—almost all the paella spots in Albufera do lunch only. We should do the lake first, then City of Arts in the afternoon. > holds, pauses

56.9–57.4 **S2** [backchannel] yeah

57.5–61.4 **S2** [failed claim] And what about Madrid? Is this just museums and walking around or—

60.6–67.9 **S1** [turn] —a mix, yeah. We'll do the Prado and Reina Sofia, but also the food markets and maybe a day trip to Toledo if everyone's up for it. > yields to S2

62.8–63.6 **S2** [backchannel] mhm

68.2–73.5 **S2** [turn] For the intercity trains—should we get some kind of Spain rail pass or just buy tickets separately? > yields to S1

73.5–81.4 **S1** [turn] I checked actually, and with only two train rides, it's cheaper to get separate tickets. The passes work out if you're taking like five or more in a week. > holds, pauses

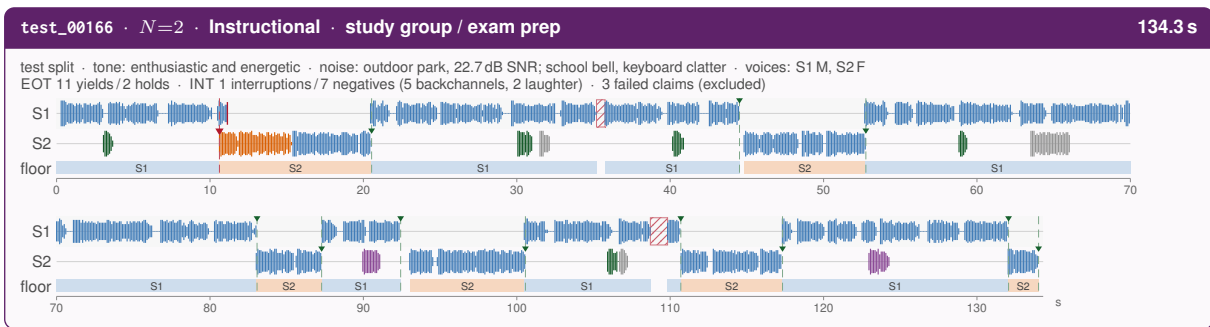
79.5–80.6 **S2** [backchannel] Makes sense

81.9–84.6 **S1** [turn] Any other must-sees for you? Or stuff you want to skip? > yields to S2

84.6–92.5 **S2** [turn] I'm good, as long as I get churros in Madrid and at least one morning by the sea in Barcelona. Oh, and no touristy flamenco shows, please. > yields to S1

92.5–99.2 **S1** [turn] Haha, got it. I'll skip those on the itinerary. Seriously, thanks for the input, this'll help me finalize the bookings tonight. > yields, end

96.6–97.5 **S2** [backchannel] Perfect



Transcript of test_00166

start-end (s) speaker [ground-truth label] text > outcome

0.3–11.1 **S1** [turn] Alright, so let's dig into photosynthesis today, since it showed up like three times on last year's exam. The big thing to remember is the two main stages: light-dependent reactions and the Calvin cycle. > cut off

3.1–3.7 **S2** [backchannel] Right.

10.6–15.3 **S2** [interruption] So, just to clarify, the light-dependent reactions happen in the thylakoid membranes, don't they? > takes floor from S1

15.4–20.5 **S2** [turn] And that's where the chlorophyll absorbs the sunlight and makes ATP and NADPH, is that right? > yields to S1

20.5–35.2 **S1** [turn] Exactly, yes! The light energy is absorbed by chlorophyll, energizing electrons which move through the electron transport chain. That makes ATP and NADPH, both of which are needed for the Calvin cycle. > holds, pauses

30.1–31.1 **S2** [backchannel] Oh, okay.

31.5–32.2 **S2** [failed claim] And—

35.8–44.5 **S1** [turn] —what's really key is remembering that oxygen gets released here, as a byproduct when water is split. That's often the part folks mix up. > yields to S2

40.2–40.9 **S2** [backchannel] Got it.

44.8–52.8 **S2** [turn] So you said ATP and NADPH go to the Calvin cycle. What exactly happens there? I always get confused with the carbon fixation bit. > yields to S1

52.7–69.5 **S1** [turn] Yeah, that's super common! In the Calvin cycle—also called the light-independent reactions—carbon dioxide from the air is fixed into glucose. That takes place in the stroma, using the energy from ATP and the electrons from NADPH.

58.8–59.4 **S2** [backchannel] Mm-hm.

63.5–66.1 **S2** [failed claim] And this is where the enzyme Rubisco comes in, right?

69.6–83.1 **S1** [turn] —yeah, exactly! Rubisco helps attach CO₂ to a five-carbon sugar, which kicks off the whole process. Then, after several steps, you get G3P, which plants use to make glucose and other things. > yields to S2

83.0–87.3 **S2** [turn] So basically, light reactions make the energy, Calvin cycle builds the sugar. > yields to S1

87.3–92.4 **S1** [turn] That's totally it! If you remember that split, you're golden on a diagram question. > yields to S2

90.0–91.1 **S2** [laughter] Haha, okay.

93.0–100.6 **S2** [turn] So, um, can I just double-check—if a question asks about where oxygen comes from, it's from the light-dependent reaction, not from splitting CO₂, right? > yields to S1

100.5–108.7 **S1** [turn] Yes, 100%! Plants split water, not carbon dioxide, to make oxygen. That's actually a classic trick question. > holds, pauses

105.9–106.6 **S2** [backchannel] Nice.

106.7–107.3 **S2** [failed claim] Wait—

109.8–110.7 **S1** [turn] —yeah go ahead! > yields to S2

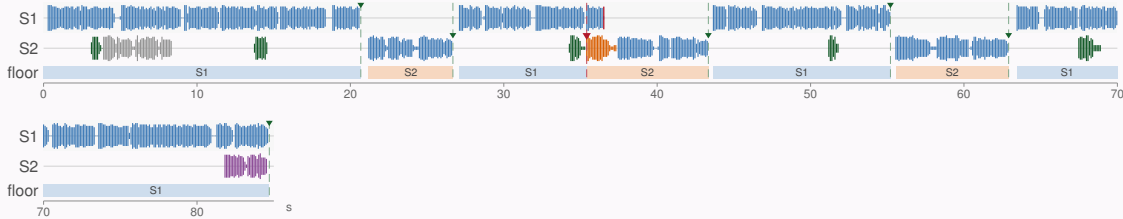
110.7–117.3 **S2** [turn] Sorry, I just thought of one more thing. Uh, do you have a good way to remember the equations for both stages? I always blank on the test. > yields to S1

117.3–132.1 **S1** [turn] Honestly? I just remember 'CO₂ in, sugar out' for the Calvin cycle, and 'water in, oxygen out' for the light-dependent reactions. But if you want the whole big equation, we can try making up a silly mnemonic right now! > yields to S2

123.0–124.3 **S2** [laughter] Haha, yes please!

132.0–134.0 **S2** [turn] That would save my life on this test. > yields, end

test split · tone: enthusiastic and energetic · noise: office hum, 35.8 dB SNR · voices: S1 F, S2 M
EOT 6 yields/0 holds · INT 1 interruptions/6 negatives (5 backchannels, 1 laughter) · 1 failed claims (excluded)

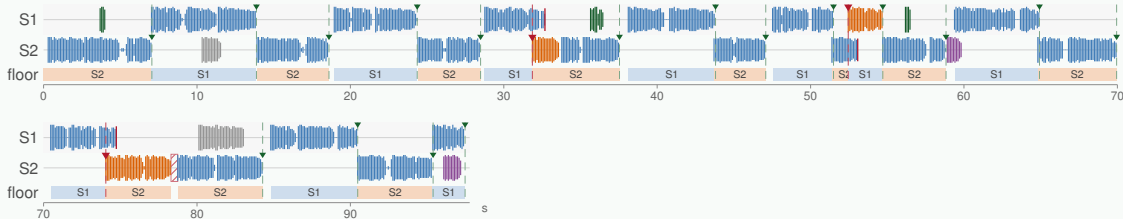


Transcript of test_00182

start-end (s) speaker [ground-truth label] text ▷ outcome

- 0.3–11.5 **S1** [turn] Alright everyone, so tonight I thought we could break down how to really analyze theme in the novel. The first thing I always recommend is to look for patterns in the characters' actions—like, what do we see Naomi doing again and again?
- 3.1–3.8 **S2** [backchannel] Uh-huh.
- 3.9–8.4 **S2** [failed claim] So are you saying—wait, do you mean like, tracking when she goes to the market, that kind of thing?
- 11.6–20.7 **S1** [turn] —kind of! It can be something as literal as her trips, but I was thinking more about, like, her tendency to help people even when it's inconvenient for her. That says something about her values, right? ▷ yields to S2
- 13.8–14.6 **S2** [backchannel] Oh, right.
- 21.2–26.7 **S2** [turn] So is that where the, uh, theme of self-sacrifice comes in? I think I spotted that but wasn't totally sure. ▷ yields to S1
- 27.1–36.5 **S1** [turn] Exactly! Yes, you've got it. And then what you can do is look for moments where that trait is challenged. Like, when her brother asks her for money, she hesitates for the first time—that's the author testing her core belief. ▷ cut off
- 34.3–35.3 **S2** [backchannel] Uh-huh, yeah.
- 35.4–37.4 **S2** [interruption] Wait, can I just jump in— ▷ takes floor from S1
- 37.5–43.3 **S2** [turn] So are we supposed to always agree on one theme? Last time, I think I saw people sort of debate different ones. ▷ yields to S1
- 43.7–55.2 **S1** [turn] No, not at all! There can definitely be more than one theme. Sometimes authors weave in two or three big ones, and it's totally normal for readers to notice different things. What matters is backing up your idea with stuff from the text. ▷ yields to S2
- 51.2–51.9 **S2** [backchannel] Right.
- 55.6–62.9 **S2** [turn] So, uh, what about that scene at the end by the river? I felt like that was almost a separate idea—about forgiveness or, or letting go, maybe. ▷ yields to S1
- 63.5–75.6 **S1** [turn] Oh I love that you brought that up! Yes, the river scene is a perfect example. You could make a case for forgiveness being a whole theme on its own. If you notice, the author uses all this water imagery—it's all washing away, old wounds and all.
- 67.5–68.9 **S2** [backchannel] Yeah! Exactly.
- 75.7–84.7 **S1** [turn] So, for next time, maybe as you read try jotting down little notes when you spot those repeated moments or symbols. The more you practice, the more these themes sort of, well, jump out at you. ▷ yields, end
- 81.8–84.6 **S2** [laughter] Haha, alright, I'll try. I appreciate this, really.

dev split · tone: curious and inquisitive · noise: school hallway, 33.5 dB SNR; keyboard clatter · voices: S1 F, S2 M
EOT 17 yields/1 holds · INT 3 interruptions/5 negatives (3 backchannels, 2 laughter) · 2 failed claims (excluded)



Transcript of dev_00005

start-end (s) speaker [ground-truth label] text ▷ outcome

- 0.3–7.1 **S2** [turn] Alright so, if we're looking at the second week of June, flights to Lisbon are actually the cheapest then. Should we just build the trip around that? ▷ yields to S1
- 3.7–4.1 **S1** [backchannel] Mm-hm
- 7.0–11.3 **S1** [turn] Wait, but have you checked what the weather's like? Because I remember it raining basically—
- 10.3–11.6 **S2** [failed claim] It hardly rains in—
- 11.4–13.9 **S1** [turn] —all the time when I went in early June last time. ▷ yields to S2
- 13.9–18.6 **S2** [turn] Right, but I think that's rare, like, it usually clears up by then? Um, I can double-check that though. ▷ yields to S1
- 18.9–24.4 **S1** [turn] Yeah, maybe just check. Also, what about Porto? Like, would we train up there or are we staying only in Lisbon? ▷ yields to S2
- 24.4–28.5 **S2** [turn] No, I really want to see Porto too. The train's not that long, right? Three hours or so? ▷ yields to S1
- 28.8–32.7 **S1** [turn] Yeah, I think it's about three. But what about the Douro Valley though? Is it on the way? Or— ▷ cut off
- 31.9–33.6 **S2** [interruption] We could stop there overnight maybe— ▷ takes floor from S1
- 33.8–37.5 **S2** [turn] Like, break up the train trip? That could actually be amazing, do some wine tasting. ▷ yields to S1
- 35.7–36.5 **S1** [backchannel] Exactly!
- 38.1–43.8 **S1** [turn] Only, uh, remember you said you wanted to see the coast? Would we have time for, like, Nazaré or some surfing town? ▷ yields to S2
- 43.7–47.1 **S2** [turn] Oh, yeah, I totally forgot! Nazaré's a little out of the way though, isn't it? ▷ yields to S1
- 47.5–51.5 **S1** [turn] It is, but if we rent a car, it's sort of doable. Or is that too ambitious? ▷ yields to S2
- 51.4–53.1 **S2** [turn] Honestly, I like ambitious. But— ▷ cut off
- 52.5–54.7 **S1** [interruption] But we only have, like, seven days, right? ▷ takes floor from S2 ▷ yields to S2
- 54.7–58.8 **S2** [turn] Yeah, seven, unless I can squeeze an extra vacation day. Maybe? I can ask. ▷ yields to S1
- 56.2–56.6 **S1** [backchannel] Ask.
- 58.9–59.9 **S2** [laughter] Haha, noted.
- 59.4–64.9 **S1** [turn] So, say we start in Lisbon, then maybe a night in the Douro, two in Porto, and then... do you want to end at the beach? ▷ yields to S2
- 64.8–70.0 **S2** [turn] That could work, yeah. But which beach? Like, I'm not a strong surfer, I just want to see the cliffs and relax. ▷ yields to S1
- 70.5–74.7 **S1** [turn] Ericeira maybe? It's closer to Lisbon for the last night. But I'm worried it's crowded. ▷ cut off

74.1–78.3 **S2**[interruption] Yeah, maybe too touristy? Wait, what about Comporta? Is that the place with the sand dunes? **> takes floor from S1 > holds, pauses**

78.7–81.2 **S2**[turn] It's south of Lisbon, so we'd have to loop around, but—

80.1–83.1 **S1**[failed claim] Yeah, Comporta! But that's, like, a fancy spot now, right?

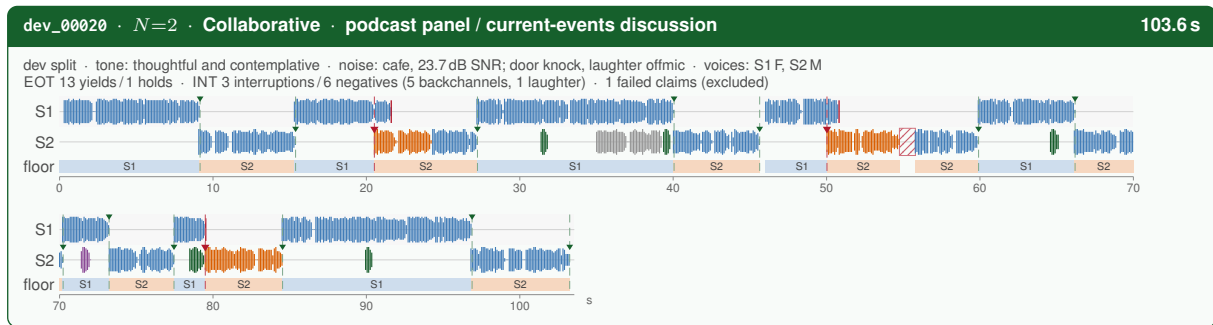
81.3–84.3 **S2**[turn] —yeah, it's gotten bougie. I don't know if that's our vibe, honestly. **> yields to S1**

84.8–90.5 **S1**[turn] No, but we can check prices. We don't need to decide everything right now, but let's... let's make a rough sketch? **> yields to S2**

90.5–95.4 **S2**[turn] Totally agree. I like where we're headed though. Should I make a shared doc and you can poke holes in my ideas? **> yields to S1**

95.4–97.5 **S1**[turn] Definitely. That's my specialty, right? **> yields, end**

96.1–97.2 **S2**[laughter] Haha, exactly.



Transcript of dev_00020

start—end (s) speaker [ground-truth label] text **> outcome**

0.3–9.2 **S1**[turn] Alright, welcome back everyone. Today, I think we really need to talk about the new climate policy proposal that dropped yesterday—there's a lot to unpack. **> yields to S2**

9.1–15.4 **S2**[turn] Yeah, completely. This is, I mean, the first time we've seen such aggressive timelines from any administration in years. **> yields to S1**

15.3–21.6 **S1**[turn] Right, and some of those targets—2030 for carbon neutrality in power generation, that's ambitious. **> cut off**

20.5–24.2 **S2**[interruption] But is it even realistic? I mean, look at the grid infrastructure— **> takes floor from S1**

24.3–27.2 **S2**[turn] —like, the investment hasn't matched the rhetoric, you know? **> yields to S1**

27.2–33.2 **S1**[turn] Absolutely, and to your point, the funding mechanisms they're proposing don't really address that bottleneck.

31.4–31.9 **S2**[backchannel] mm-hm

33.3–38.2 **S1**[turn] There's also a lot of, um, emphasis on innovation, but not as much on regulation.

35.0–39.3 **S2**[failed claim] But do we even need more regulation or just enforcement of what's already—

38.3–40.1 **S1**[turn] —on the books? That's the question, right? **> yields to S2**

39.4–39.9 **S2**[backchannel] yeah

40.0–45.6 **S2**[turn] I mean, if you just take methane leaks, for example, we've had standards, but barely any compliance checks. **> yields to S1**

46.0–50.8 **S1**[turn] Exactly, enforcement's been spotty. But isn't there a genuine shift right now? Like, I feel there's more willingness across states to collaborate. **> cut off**

50.0–54.7 **S2**[interruption] But don't you think part of that is, uh, political pressure from international actors? **> takes floor from S1 > holds, pauses**

55.8–59.9 **S2**[turn] I saw that Europe basically threatened tariffs unless our standards matched. **> yields to S1**

59.9–66.2 **S1**[turn] Absolutely, that's a huge factor. We're seeing climate policy morph into trade policy almost overnight. **> yields to S2**

64.6–65.2 **S2**[backchannel] right

66.2–70.2 **S2**[turn] And that's where local industry gets, well, nervous. Like, the auto sector— **> yields to S1**

70.2–73.2 **S1**[turn] Yeah, the auto sector's got a target on its back right now. **> yields to S2**

71.4–72.0 **S2**[laughter] haha

73.2–77.5 **S2**[turn] I mean, they argue it's too fast, but then, you look at what's happening in China— **> yields to S1**

77.5–79.5 **S1**[turn] They've leapfrogged us on EVs, completely. **> cut off**

78.5–79.4 **S2**[backchannel] Exactly.

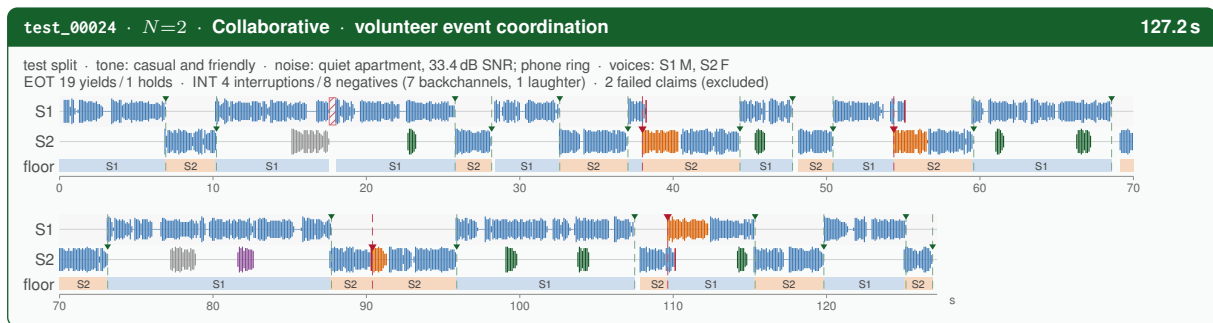
79.5–84.5 **S2**[interruption] But so, do you think this policy package is gonna move the needle, or is it just more posturing? **> takes floor from S1 > yields to S1**

84.5–92.6 **S1**[turn] I think it's a bit of both, honestly. The ambition will set a tone, but without serious funding and buy-in, it's gonna stall—like so many before.

89.9–90.4 **S2**[backchannel] mm-hm

92.7–96.9 **S1**[turn] But, it does matter symbolically, and I think public momentum is still building. **> yields to S2**

96.8–103.3 **S2**[turn] For sure. And I guess the real test is if, two years from now, we're even talking about implementation instead of goals. **> yields, end**



Transcript of test_00024

start—end (s) speaker [ground-truth label] text **> outcome**

0.3–6.9 **S1**[turn] Alright, uh, so thanks for hopping on, Mia. I wanted to talk through the logistics for Saturday's tree planting thing. **> yields to S2**

6.9–10.2 **S2**[turn] Of course! I'm excited, honestly. Did you get enough folks to sign up? **> yields to S1**

10.2–17.6 **S1**[turn] Yeah! We've got like fifteen volunteers confirmed so far, but I'm a little worried about the tool situation. Last spring— **> holds, pauses**

15.1–17.6 **S2**[failed claim] Oh, wait, do we still have shovels from Parks, or—

18.0–25.8 **S1**[turn] —right, that's the thing, they only have nine shovels they can lend us. So I was thinking maybe we ask people to bring extras? **> yields to S2**

22.7–23.3 **S2**[backchannel] mm-hm

25.8–28.2 **S2** [turn] I can send a reminder email for that, yeah. > yields to S1

28.4–32.6 **S1** [turn] Perfect. And uh, snacks! You mentioned maybe doing granola bars? > yields to S2

32.6–37.0 **S2** [turn] Haha, always snacks. I can totally pick some up, and I'll bring a big jug of water too. > yields to S1

37.1–38.2 **S1** [turn] Awesome! Oh, timing— > cut off

38.0–40.4 **S2** [interruption] Oh, right, should we start at ten or is nine better? > takes floor from S1

40.5–44.4 **S2** [turn] Some folks mentioned their kids had soccer, so ten might be easier for them. > yields to S1

44.4–47.8 **S1** [turn] Good call, yeah. Ten works. I'll update the event flyer. > yields to S2

45.4–46.0 **S2** [backchannel] yeah

48.2–50.4 **S2** [turn] What about rain? Was there a backup plan or...? > yields to S1

50.5–55.1 **S1** [turn] So, I was thinking we could push to Sunday if it really pours. Or, if it's just a drizzle— > cut off

54.4–56.6 **S2** [interruption] People'll come anyway, I bet. Last year— > takes floor from S1

56.7–59.6 **S2** [turn] Basically everyone just wore raincoats and kept digging. > yields to S1

59.5–62.4 **S1** [turn] Haha, yeah, volunteers are tougher than the weather.

61.0–61.6 **S2** [backchannel] right

62.5–68.6 **S1** [turn] But you know, I'll set up a group text so we can send a last-minute update if the forecast looks bad. > yields to S2

66.3–67.2 **S2** [backchannel] good idea

69.1–73.1 **S2** [turn] Speaking of that, do you know if the city's sending anyone to talk about the new park plans? > yields to S1

73.1–79.1 **S1** [turn] Um, I emailed them, but I haven't heard back yet. Last time they just sent Steve with his maps—

77.2–78.9 **S2** [failed claim] He's got the best stories though.

79.2–82.8 **S1** [turn] —right? He always gets people hyped about pollinator gardens.

81.6–82.7 **S2** [laughter] haha totally

83.0–87.7 **S1** [turn] So, uh, do you wanna handle the sign-in sheet? I can drop it off Friday if you want. > yields to S2

87.6–90.3 **S2** [turn] Sure, I don't mind. I'll print extras just in case. > cut off

90.4–91.4 **S2** [interruption] oh! And— > takes floor from S1

91.5–95.9 **S2** [turn] Um, do we have enough gloves? I think the bin was looking empty when we packed up last time. > yields to S1

95.8–100.8 **S1** [turn] Uh, you're right, I should... I'll check what's left in storage and pick up a pack if we're short.

99.1–99.9 **S2** [backchannel] awesome

100.9–107.5 **S1** [turn] Anything else? Oh, wait—parking! Folks should park on Elm, not in the rec lot. I'll put it in the directions. > yields to S2

103.8–104.6 **S2** [backchannel] good call

107.9–110.1 **S2** [turn] I think that covers most of it? Unless— > cut off

109.6–112.3 **S1** [interruption] Just double-checking everyone's signed the new waiver. > takes floor from S2

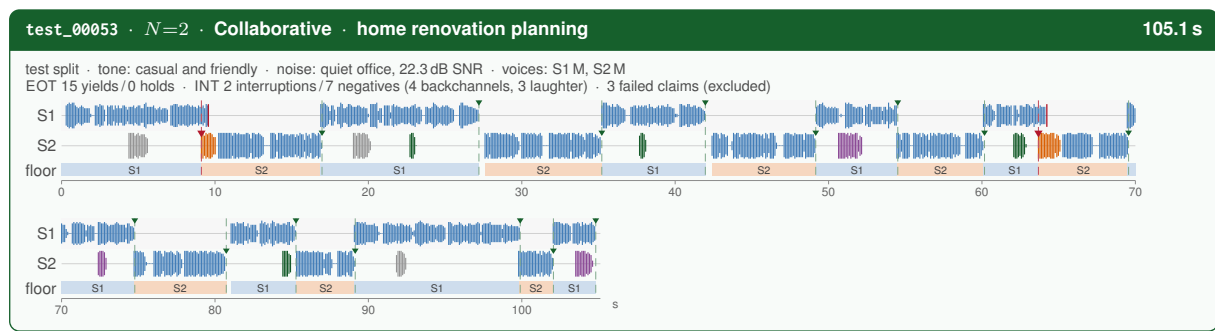
112.4–115.3 **S1** [turn] I'll attach it to your reminder email if that works. > yields to S2

114.2–114.9 **S2** [backchannel] perfect

115.3–119.8 **S2** [turn] This is shaping up to be so much fun. Thanks for pulling all the pieces together, Alex. > yields to S1

119.9–125.2 **S1** [turn] Haha, right back at you, Mia. I think we're set—unless you think of something wild last minute? > yields to S2

125.1–126.9 **S2** [turn] No promises! See you Saturday. > yields, end



Transcript of test_00053

start—end (s) speaker [ground-truth label] text > outcome

0.3–7.0 **S1** [turn] Alright, so, for the living room—hear me out—I really think knocking out that wall by the kitchen would just open everything up, you know?

4.4–5.7 **S2** [failed claim] Yeah, I see what you mean.

7.1–9.6 **S1** [turn] We could do one of those, uh, island counters instead. Maybe with some butcher block— > cut off

9.1–10.1 **S2** [interruption] But the thing is— > takes floor from S1

10.2–17.0 **S2** [turn] —if we knock out that wall, won't we lose all the cabinet space? I mean, I get the open feel, but where do we put, like, half our stuff? > yields to S1

16.9–20.0 **S1** [turn] Oh, totally fair. We, uh, could do ceiling-hung shelves or—

19.0–20.2 **S2** [failed claim] Like open shelving?

20.1–27.2 **S1** [turn] —yeah, exactly. I mean, it'd look a little, you know, 'Pinterest-y,' but we could keep stuff we use a lot right there. Rest goes in the pantry. > yields to S2

22.7–23.1 **S2** [backchannel] mm-hm

27.6–35.2 **S2** [turn] I just—man, I hate dusting open shelves. Everything gets so gross, right? But it does look good. Maybe like, partial open? Some glass doors? > yields to S1

35.2–42.0 **S1** [turn] Glass doors! Now we're talkin'. We could, uh, do those on top and closed cabinets below. That way, it's not all out in the open. > yields to S2

37.7–38.1 **S2** [backchannel] yeah

42.4–49.2 **S2** [turn] So, about the floors—are we actually gonna try refinishing, or just rip 'em out? 'Cause those pine boards are pretty beaten up. > yields to S1

49.2–54.5 **S1** [turn] If it were up to me, I'd keep 'em. Adds character, y'know? But, uh, I know you're not a fan of the squeaks. > yields to S2

50.6–52.2 **S2** [laughter] haha, no, they drive me nuts.

54.5–60.2 **S2** [turn] What if—okay, what if we refinish but put rugs down in the main walkways? That hides some of it, right? > yields to S1

60.1–64.2 **S1** [turn] Yeah, could work. Plus, we could try that, uh, thing where you sprinkle talc powder in the cracks. Supposed to quiet the boards down? > cut off

62.1–62.9 **S2** [backchannel] Really? Huh.

63.7–65.1 **S2** [interruption] Okay, so, for the bathroom— > takes floor from S1

65.3–69.6 **S2** [turn] Do you still want to move the sink under the window? Or was that just, like, a passing thought? > yields to S1

69.5–74.8 **S1** [turn] Nah, still want it. Morning light's perfect there, man. Just makes shaving feel less, I dunno, prison-y. > yields to S2

72.4–73.0 **S2** [laughter] haha

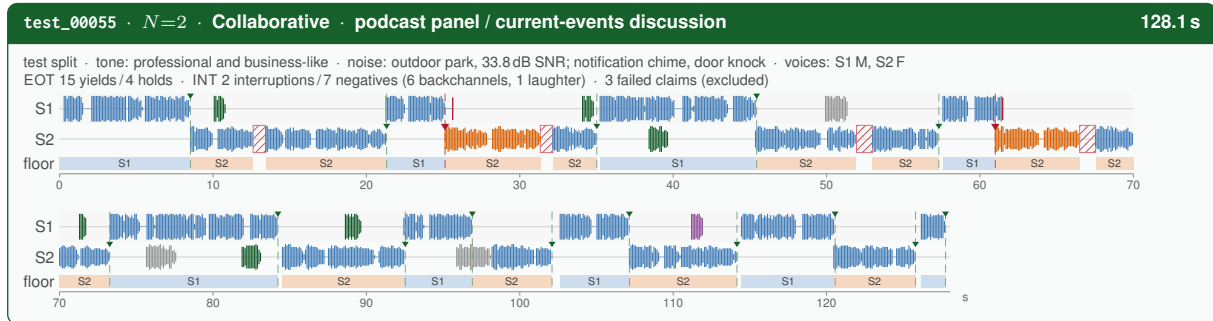
74.7–80.7 **S2** [turn] Alright, I'm in. But if the old plumbing gives us trouble, you're helping me bail water out of the crawl space again. > yields to S1

81.1–85.3 **S1** [turn] Deal. Hey, what about the front porch? That railing's basically chewing gum at this point. > yields to S2

84.4–85.0 **S2** [backchannel] mm, yeah

85.3–89.1 **S2** [turn] Can we go composite, or do you wanna stick with wood? I know you like the look— > yields to S1

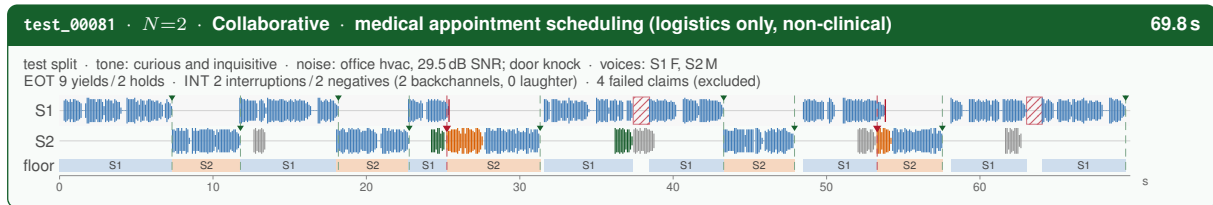
89.1–95.1 **S1** [turn] Composite's fine, honestly. No more splinters in my palm, please. Let's just match the color close to the original.
 91.8–92.5 **S2** [failed claim] Sounds—
 95.2–99.9 **S1** [turn] —good. And hey, if we finish all this, I get dibs on picking the paint color for the entryway. $\triangleright$ yields to S2
 99.8–102.1 **S2** [turn] Only if you promise not to pick orange again. $\triangleright$ yields to S1
 102.1–104.8 **S1** [turn] No promises, man. But I'll try to restrain myself. $\triangleright$ yields, end
 103.5–104.7 **S2** [laughter] haha, alright, deal.



Transcript of test_00055

start–end (s) speaker [ground-truth label] text $\triangleright$ outcome

0.3–8.5 **S1** [turn] Welcome back to the panel. Today we're diving into the latest on remote work policy shifts as companies start calling employees back to the office. $\triangleright$ yields to S2
 8.5–12.6 **S2** [turn] Right, and it's moving quickly. I mean, Google's announcement this week was— $\triangleright$ holds, pauses
 10.1–10.8 **S1** [backchannel] Big. Yeah.
 13.4–21.3 **S2** [turn] —frankly, a bellwether. They're shifting to a three-two model, which means employees are required in the office at least three days a week, no exceptions. $\triangleright$ yields to S1
 21.3–25.6 **S1** [turn] But do you think that'll stick? Because Meta tried a similar approach in the spring. Morale tanked, productivity metrics didn't budge— $\triangleright$ cut off
 25.1–31.3 **S2** [interruption] —but Meta's rollout was abrupt, no clear comms. Google's staggering their return dates, it's a more phased strategy. $\triangleright$ takes floor from S1 $\triangleright$ holds, pauses
 32.2–35.0 **S2** [turn] And they're offering childcare perks, which Meta didn't. $\triangleright$ yields to S1
 34.1–34.8 **S1** [backchannel] Good point.
 35.2–45.4 **S1** [turn] I've heard from managers that the hybrid schedule works only if teams can coordinate their days. Otherwise, half the collaboration value gets lost. You end up on more Zooms. $\triangleright$ yields to S2
 38.4–39.7 **S2** [backchannel] Yeah, exactly.
 45.4–51.9 **S2** [turn] But you know, some teams are leveraging that by splitting focus days and meeting days. The data from Slack's Future Forum— $\triangleright$ holds, pauses
 49.9–51.4 **S1** [failed claim] Oh, the pulse surveys, yeah.
 53.0–57.3 **S2** [turn] —shows higher satisfaction when there's predictability, not just flexibility. $\triangleright$ yields to S1
 57.6–61.5 **S1** [turn] But there's still a leadership gap, right? The C-suite is mostly back, but middle managers are caught in between and— $\triangleright$ cut off
 61.0–66.5 **S2** [interruption] —that's exactly it, they're under pressure from both sides. I spoke to a director at Salesforce who said— $\triangleright$ takes floor from S1 $\triangleright$ holds, pauses
 67.5–73.3 **S2** [turn] She's managing expectations from execs demanding visibility and employees demanding autonomy. $\triangleright$ yields to S1
 71.3–71.8 **S1** [backchannel] Right.
 73.3–79.6 **S1** [turn] And nobody wants to be first to push back. I mean, we saw what happened at Amazon last month, with that open letter—
 75.7–77.6 **S2** [failed claim] —the employee walkouts, yes.
 79.8–84.2 **S1** [turn] —and the comms team scrambling to get ahead of it. It's a huge reputational risk. $\triangleright$ yields to S2
 81.9–83.1 **S2** [backchannel] Completely agree.
 84.5–92.5 **S2** [turn] But can I challenge one point? The productivity numbers. I think the issue is, most companies don't know what to measure yet. The metrics are lagging reality. $\triangleright$ yields to S1
 88.6–89.7 **S1** [backchannel] Interesting, yes.
 92.4–96.9 **S1** [turn] So you're saying output is up, but the KPIs are stuck in a 2019 mindset. $\triangleright$ yields to S2
 95.9–98.1 **S2** [failed claim] Exactly. Especially for knowledge work.
 98.2–102.1 **S2** [turn] And yet, leaders still gravitate to butts-in-seats as a comfort metric. $\triangleright$ yields to S1
 102.6–107.1 **S1** [turn] You know, that's not changing anytime soon. I think it's partly generational, honestly. $\triangleright$ yields to S2
 107.1–114.2 **S2** [turn] But, well, Gen Z is pushing back hard. Look at TikTok—there's a whole subculture making fun of badge swipes and office mandates. $\triangleright$ yields to S1
 111.2–112.0 **S1** [laughter] Haha, yeah.
 114.5–120.6 **S1** [turn] So, looking ahead, do you actually think the three-two model is here to stay? Or are we just in another holding pattern? $\triangleright$ yields to S2
 120.5–125.8 **S2** [turn] Honestly? I think it'll keep morphing—a year from now, there'll be something else. But we're learning, fast. $\triangleright$ yields to S1
 126.2–127.8 **S1** [turn] Fair enough. Let's keep watching. $\triangleright$ yields, end

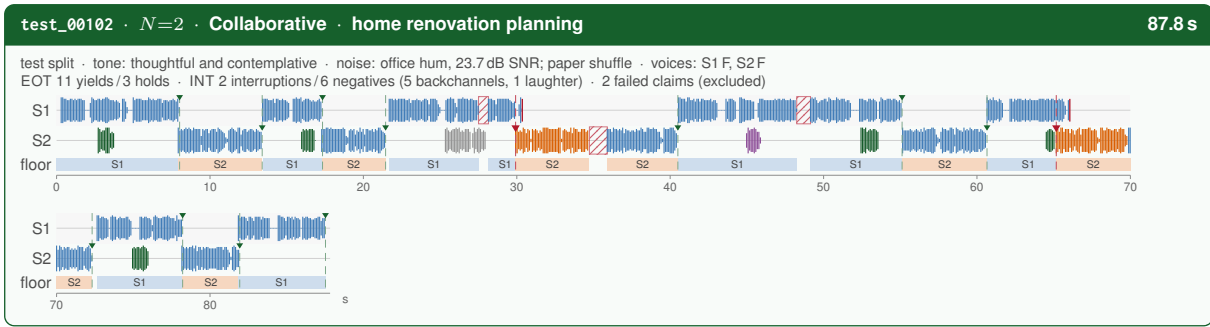


Transcript of test_00081

start–end (s) speaker [ground-truth label] text $\triangleright$ outcome

0.3–7.3 **S1** [turn] So, to clarify, we're looking for the soonest possible appointment but something after four p.m. works best, right? $\triangleright$ yields to S2
 7.4–11.8 **S2** [turn] Yeah, uh, that's right. I— I can do a little earlier if there's nothing after four though. $\triangleright$ yields to S1
 11.8–13.5 **S1** [turn] Let me just check the system, um—
 12.7–13.5 **S2** [failed claim] Wait, do we know—
 13.6–18.2 **S1** [turn] —whether it's in-person only or if telehealth is okay? That changes what's open. $\triangleright$ yields to S2

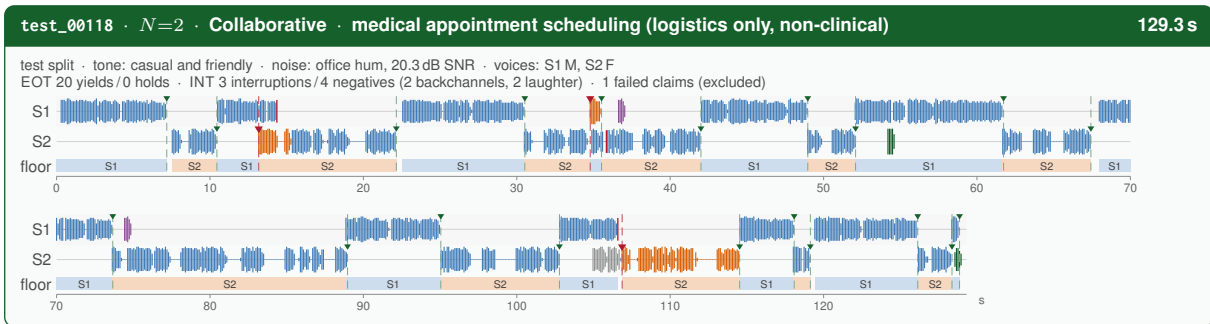
18.1–22.8 **S2** [turn] Oh, uh, telehealth would actually be better for me. Sorry, I should've said that sooner. ▷ yields to S1
 22.7–25.4 **S1** [turn] No, that helps. So telehealth after four— ▷ cut off
 24.2–25.2 **S2** [backchannel] Yeah, exactly.
 25.2–27.6 **S2** [interruption] Do you see anything like, uh, Thursday or Friday? ▷ takes floor from S1
 27.7–31.3 **S2** [turn] Because I tend to get off work early those days if that helps at all. ▷ yields to S1
 31.6–37.4 **S1** [turn] Got it, Thursday or Friday, after four, telehealth. I'm seeing a slot Thursday at 4:30? ▷ holds, pauses
 36.2–37.4 **S2** [backchannel] Oh, 4:30's good.
 37.4–38.8 **S2** [failed claim] Wait, is that with Dr. Lin?
 38.5–43.3 **S1** [turn] —No, that's with Dr. Patel. Would that still work, or do you need Lin specifically? ▷ yields to S2
 43.3–47.9 **S2** [turn] I've only seen Dr. Lin before, but, um... is Patel okay for follow-ups? ▷ yields to S1
 48.5–53.8 **S1** [turn] She's fine for follow-ups, yeah. Actually most of the telehealth ones are with her right now. ▷ cut off
 52.1–53.2 **S2** [failed claim] Alright, that works for me.
 53.3–54.2 **S2** [interruption] But is it possible— ▷ takes floor from S1
 54.3–57.5 **S2** [turn] Sorry, is it possible to switch if Lin opens up last minute? ▷ yields to S1
 58.1–63.0 **S1** [turn] Absolutely. You can swap if her schedule frees up, just let us know before the day. ▷ holds, pauses
 61.7–62.8 **S2** [failed claim] Great, thanks so much.
 64.1–69.5 **S1** [turn] No problem—so Thursday, 4:30, Dr. Patel, telehealth. I'll confirm it now. ▷ yields, end



Transcript of test_00102

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.0 **S1** [turn] Okay, so picture this—what if we turn the old dining room into, like, a reading nook instead of just another boring eating area? ▷ yields to S2
 2.7–3.8 **S2** [backchannel] Oh I love—
 7.9–13.4 **S2** [turn] —that idea, but I worry it might feel cramped with the bookshelves? You know, since that wall is kinda short. ▷ yields to S1
 13.4–17.3 **S1** [turn] Yeah but if we do floating shelves, it wouldn't eat up so much space, right? ▷ yields to S2
 16.0–16.9 **S2** [backchannel] Mmm, maybe—
 17.3–21.5 **S2** [turn] But what about that little window? Do you think the sun would bleach all the covers? ▷ yields to S1
 21.7–27.5 **S1** [turn] Oh! We could put up those sheer curtains, you know, just to take the edge off? Not, like, full blackout or anything. ▷ holds, pauses
 25.4–27.1 **S2** [failed claim] Yeah, sheer—yeah, that might work.
 27.1–28.0 **S2** [failed claim] But wait—
 28.2–30.4 **S1** [turn] —and maybe, oh! A window seat! That would make it super cozy— ▷ cut off
 29.9–34.7 **S2** [interruption] —yes, exactly! Window seat, and then maybe, um, like, cushions that double as storage? ▷ takes floor from S1 ▷ holds, pauses
 35.9–40.5 **S2** [turn] Because, honestly, we have way too many board games floating around. They need a home. ▷ yields to S1
 40.5–48.2 **S1** [turn] Haha, yes, the Great Board Game Invasion! We could, like, hide them under the seat, never to be seen until someone begs for Scrabble. ▷ holds, pauses
 45.0–45.9 **S2** [laughter] Haha right
 49.1–55.1 **S1** [turn] But, um, do you think the colors should match the living room? Or should we try, like, something totally different in there? ▷ yields to S2
 52.4–53.6 **S2** [backchannel] Oh, that's tricky—
 55.2–60.7 **S2** [turn] I kind of want it to feel separate, but not like... jarringly so? Maybe just echo some little accents? ▷ yields to S1
 60.7–66.1 **S1** [turn] Echo! I like that. Like, keep the same yellow pillows but swap out the blue rug for, um, maybe something green? ▷ cut off
 64.5–65.1 **S2** [backchannel] Mm-hm
 65.2–69.8 **S2** [interruption] Wait—actually, can we keep that plant in there, too? I feel like it needs a little life. ▷ takes floor from S1
 69.9–72.3 **S2** [turn] Just, like, one big leafy guy in the corner? ▷ yields to S1
 72.6–78.2 **S1** [turn] Absolutely! It'll be the plant's new throne. But seriously, we might need to measure everything first— ▷ yields to S2
 75.0–76.0 **S2** [backchannel] Yeah, totally.
 78.2–81.9 **S2** [turn] Otherwise we'll end up with a shelf behind the door again and... yeah. ▷ yields to S1
 81.9–87.5 **S1** [turn] Haha, nope, not repeating that disaster. Okay, measuring first, then window seat dreams after. ▷ yields, end



Transcript of test_00118

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–7.2 **S1** [turn] Hey, so I was looking at the online thing for Dr. Patel, and, um, there's literally nothing open for like the next three weeks. ▷ yields to S2

7.5–10.4 **S2** [turn] Shocking. The universe conspires again. ▷ yields to S1

10.5–14.4 **S1** [turn] Yeah, right? I mean, I can do afternoons if they ever free up, but mornings are pretty much— ▷ cut off

13.2–15.3 **S2** [interruption] —totally impossible? Same. ▷ takes floor from S1

15.4–22.2 **S2** [turn] Like, if they offer me a 7:40 appointment again, I will simply cease to exist. ▷ yields to S1

22.6–30.5 **S1** [turn] Yeah, no way, I'm not even vertical before eight. Did you try calling their front desk? Sometimes they magically have slots that don't show online. ▷ yields to S2

30.5–35.8 **S2** [turn] Tried. Got put on hold for, no exaggeration, ten minutes, and then they just— ▷ cut off

34.8–35.5 **S1** [interruption] —hung up on you? ▷ takes floor from S2 ▷ yields to S2

35.9–38.9 **S2** [turn] Yeah. Like, a power move. Hardcore.

36.6–37.1 **S1** [laughter] haha

38.9–42.0 **S2** [turn] So I guess we're just fighting bots for scraps. ▷ yields to S1

42.0–49.0 **S1** [turn] Right, it's like concert tickets. Okay, what if we both keep refreshing and text if we see something? Fastest one grabs it. ▷ yields to S2

49.0–52.1 **S2** [turn] Deal. You know, 'cause I live for adrenaline. ▷ yields to S1

52.1–55.4 **S1** [turn] Exactly, nothing like medical appointment roulette before coffee.

54.2–54.7 **S2** [backchannel] uh-huh

55.5–61.7 **S1** [turn] What if we ask for a waitlist? Like, if someone cancels, maybe they'll call us before some rando snags it online. ▷ yields to S2

61.7–67.4 **S2** [turn] Oh, waitlists. The mythical creatures. Do you think that really works? ▷ yields to S1

68.0–73.7 **S1** [turn] My mom swears by it. She's got, like, black belt scheduler energy. I can ask her for tips if you want. ▷ yields to S2

73.6–76.3 **S2** [turn] Please. Maybe she'll broker us a treaty.

74.4–74.9 **S1** [laughter] haha

76.4–89.0 **S2** [turn] Seriously though, if either of us gets through to an actual human, we should just book whatever. I mean, beggars, choosers, et cetera. ▷ yields to S1

88.9–95.0 **S1** [turn] Yeah, agreed, just lock it in and we'll swap if needed. Uh, do you want me to try calling again, maybe like right at nine? ▷ yields to S2

95.1–102.8 **S2** [turn] If you have the emotional stamina, be my guest. I can try their portal thing at the same time. ▷ yields to S1

102.8–106.6 **S1** [turn] Deal. And if the phone melts my brain, you'll hear me scream from across town. ▷ cut off

105.0–106.8 **S2** [failed claim] just scream into the void

106.9–114.5 **S2** [interruption] But wait—did you try that new chat bot thing they added last month? Maybe it's less cursed. ▷ takes floor from S1 ▷ yields to S1

114.6–118.1 **S1** [turn] Oh, the bot? I typed "I want an appointment" and it offered me yoga tips. ▷ yields to S2

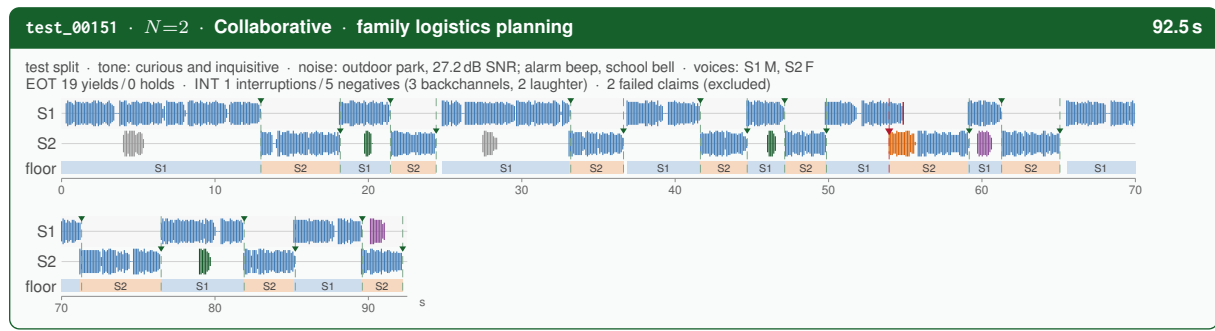
118.0–119.1 **S2** [turn] Of course it did. ▷ yields to S1

119.4–126.1 **S1** [turn] So yeah, I'll try a good ol' fashioned phone call. You keep refreshing, and whoever wins gets bragging rights for, like, a week. ▷ yields to S2

126.2–128.4 **S2** [turn] Game on. Loser buys coffee. ▷ yields to S1

128.4–128.9 **S1** [turn] You're on. ▷ yields, end

128.6–129.0 **S2** [backchannel] mm-hm



Transcript of test_00151

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.8 **S1** [turn] So, um, I was looking at the schedule for next Saturday and, I mean, have we actually figured out who's driving Emma to her swim meet?

4.0–5.4 **S2** [failed claim] Wait, I thought you were—

6.9–13.0 **S1** [turn] —I was supposed to, yeah, but then my mom asked if we could stop by in the morning, so now I'm not sure I can make it in time. ▷ yields to S2

13.0–18.2 **S2** [turn] Oh right, yeah. Um, I think I could probably do the drive, but wouldn't that overlap with grocery pick-up? ▷ yields to S1

18.1–21.4 **S1** [turn] Yeah... the Instacart slot is basically right when her warmups start. ▷ yields to S2

19.8–20.3 **S2** [backchannel] hmm

21.5–24.4 **S2** [turn] Is there any way your mom could shift her visit a little later? ▷ yields to S1

24.8–29.6 **S1** [turn] I, uh, I could ask, but she's already planned her morning around us heading there before lunch.

27.5–28.4 **S2** [failed claim] Well what if—

29.7–33.2 **S1** [turn] —but I guess if we asked pretty soon maybe she could be flexible, I don't know. ▷ yields to S2

33.1–36.6 **S2** [turn] Yeah, or, um, could Emma maybe carpool with Olivia's family this time? ▷ yields to S1

36.9–41.6 **S1** [turn] Oh, that's a good idea. I hadn't thought of the carpool. Do we know if Olivia's mom usually drives? ▷ yields to S2

41.7–44.7 **S2** [turn] Usually, yeah. And they live like five minutes from us, so it wouldn't— ▷ yields to S1

44.7–47.1 **S1** [turn] Perfect, uh, I think Emma has her number saved. ▷ yields to S2

46.0–46.6 **S2** [backchannel] mm-hm

47.2–49.9 **S2** [turn] I could text her and just ask if she's got room in the car? ▷ yields to S1

49.8–54.9 **S1** [turn] Yeah, that'd solve a lot if she's okay with it. But then we still have to figure out how Emma gets home, right? ▷ cut off

53.9–55.7 **S2** [interruption] True, I can pick her up after— ▷ takes floor from S1

55.8–59.2 **S2** [turn] —honestly after the meet I should be back from groceries, so that might work. ▷ yields to S1

59.1–61.3 **S1** [turn] Great, yeah, that's probably the simplest. ▷ yields to S2

59.7–60.7 **S2** [laughter] haha yeah

61.3–65.1 **S2** [turn] And, okay, what about dinner, are we still doing pizza night with your folks or...? ▷ yields to S1

65.5–71.3 **S1** [turn] My mom mentioned it, but she sounded kind of unsure? I think they're waiting to see if Dad's back from his conference. ▷ yields to S2

71.2–76.5 **S2** [turn] Oh, right, I forgot about that. Should we plan a backup just in case? Maybe I could make that soup Emma likes. ▷ yields to S1

76.5–81.9 **S1** [turn] Soup sounds good. Or, um, even just sandwiches if it's been a crazy day. Let's just see how things shake out. ▷ yields to S2

79.0–79.7 **S2** [backchannel] okay

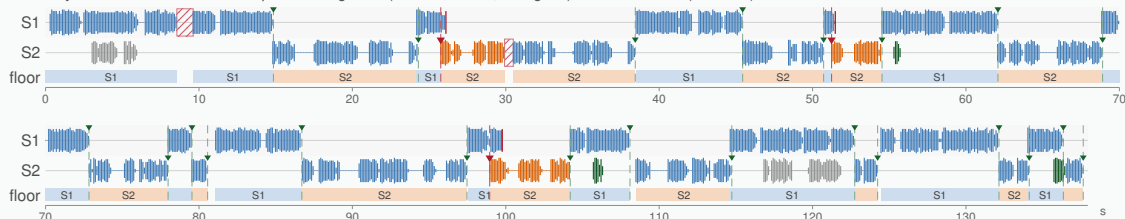
81.9–85.2 **S2** [turn] Sounds like a plan. I'll text Olivia's mom after we finish talking? ▷ yields to S1

85.1–89.6 **S1** [turn] Yeah, and I'll double-check with my mom about timings. I think that should cover most of it. ▷ yields to S2

89.5–92.2 **S2** [turn] Perfect, now if only next weekend was this easy! ▷ yields, end

90.1–91.1 **S1** [laughter] haha, right.

test split · tone: enthusiastic and energetic · noise: school hallway, 21.2 dB SNR; school bell, keyboard clatter · voices: S1 M, S2 F
EOT 23 yields / 2 holds · INT 3 interruptions / 3 negatives (3 backchannels, 0 laughter) · 2 failed claims (excluded)



Transcript of test_00174

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.6 **S1** [turn] So, um, for Saturday, I was thinking, what if we set up the tables in a U-shape instead of the usual straight rows? That way people can actually talk to each other— ▷ holds, pauses

3.0–6.0 **S2** [failed claim] You mean so they conspire more efficiently—

9.6–14.9 **S1** [turn] —well, or, uh, collaborate. I just thought it'd be friendlier than last year's, you know, cafeteria lineup. ▷ yields to S2

14.8–24.3 **S2** [turn] Yeah, I like the spirit. If someone starts an uprising for more coffee, we blame the table shape. Got it. ▷ yields to S1

24.2–26.1 **S1** [turn] Haha, that's fair. Speaking of which— ▷ cut off

25.8–29.9 **S2** [interruption] Oh coffee, yes, I was literally just about to say— ▷ takes floor from S1 ▷ holds, pauses

30.5–38.4 **S2** [turn] Last year it ran out by like, ten. Maybe this time we should double the order? Or— ▷ yields to S1

38.5–45.4 **S1** [turn] Yeah, I can call Brew Haven and see if they'll throw in a discount for a bigger batch. Uh, do we want iced coffee too? Could be warm Saturday. ▷ yields to S2

45.4–50.7 **S2** [turn] Iced coffee is fancy for us. Next you'll suggest scones. ▷ yields to S1

50.7–51.5 **S1** [turn] Actually— ▷ cut off

51.2–54.5 **S2** [interruption] Haha, no, I'm vetoing the pastry budget. ▷ takes floor from S1 ▷ yields to S1

54.5–57.2 **S1** [turn] Alright, alright, just coffee. Not even a muffin. Got it.

55.3–55.8 **S2** [backchannel] mm-hm

57.3–62.1 **S1** [turn] So, uh, what about sign-in? Do we use the sheets again, or that QR code thing you tried last month? ▷ yields to S2

62.1–68.9 **S2** [turn] I mean, the QR was ambitious. Half the people pointed their cameras at the ceiling. ▷ yields to S1

68.8–72.8 **S1** [turn] Yeah, I noticed that. Maybe just put a few clipboards out and call it modern enough. ▷ yields to S2

72.8–78.0 **S2** [turn] We could tape the pens down so they don't all mysteriously vanish... ▷ yields to S1

78.0–79.5 **S1** [turn] Pen security, I like it. ▷ yields to S2

79.5–80.6 **S2** [turn] Yeah, it's crucial. ▷ yields to S1

81.1–86.7 **S1** [turn] So, uh, the stations—can we shuffle volunteers this time? People got kind of bored at one table all morning. ▷ yields to S2

86.7–97.5 **S2** [turn] Could do like, hour-long slots? I mean unless someone develops a deep attachment to, I don't know, the water bottle table. ▷ yields to S1

97.5–99.8 **S1** [turn] Haha, I'll make a spreadsheet. Uh, actually, maybe— ▷ cut off

99.0–104.2 **S2** [interruption] Send it to everyone and watch three people actually fill it out? ▷ takes floor from S1 ▷ yields to S1

104.2–108.1 **S1** [turn] —yeah, exactly. I'll just assign people and they can swap if they threaten mutiny. ▷ yields to S2

105.7–106.3 **S2** [backchannel] Smart.

108.5–114.7 **S2** [turn] What about clean-up? Because last year everyone scattered like we yelled fire. ▷ yields to S1

114.6–119.4 **S1** [turn] Yeah, we should, uh, maybe bribe them? Like, if you stay and sweep you get extra raffle tickets or—

116.8–121.9 **S2** [failed claim] Or, we just lock the doors. Nobody leaves till it's sparkling.

119.5–122.7 **S1** [turn] —that's, um, slightly illegal? But I like the raffle idea. ▷ yields to S2

122.8–124.2 **S2** [turn] Legal shmegal. ▷ yields to S1

124.5–132.2 **S1** [turn] Alright, so recap: U-shaped tables, lots of coffee but no pastries, sign-in sheets with chained pens, volunteer shuffle, and bribed clean-up crew. ▷ yields to S2

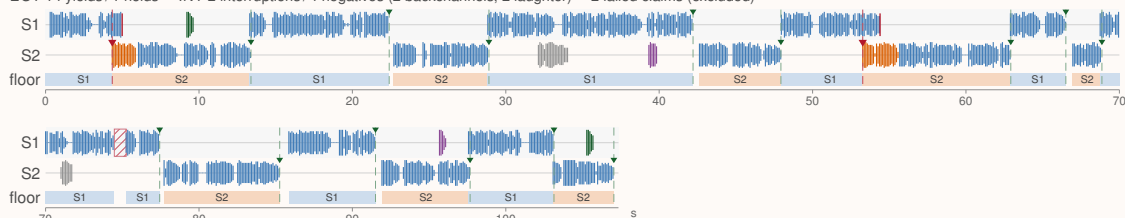
132.2–134.1 **S2** [turn] We are basically unstoppable. ▷ yields to S1

134.0–136.3 **S1** [turn] Haha, I'll send this out today. Sound good? ▷ yields to S2

135.7–136.5 **S2** [backchannel] Perfect.

136.5–137.6 **S2** [turn] See you Saturday. ▷ yields, end

dev split · tone: curious and inquisitive · noise: office hum, 32.1 dB SNR; phone ring, door knock · voices: S1 M, S2 F
EOT 14 yields / 1 holds · INT 2 interruptions / 4 negatives (2 backchannels, 2 laughter) · 2 failed claims (excluded)



Transcript of dev_00013

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–5.0 **S1** [turn] Honestly, I don't think cramming is even effective. Like, you forget it all after the test anyway. ▷ cut off

4.3–5.9 **S2** [interruption] No, that's not true at all— ▷ takes floor from S1

6.1–13.4 **S2** [turn] I mean, if you space it out you just get overwhelmed. Cramming works if you're focused, and like, it's proven that adrenaline helps memory. ▷ yields to S1

9.2–9.7 **S1** [backchannel] mm-hm

13.3–22.4 **S1** [turn] That sounds fake. No, the studies literally say spaced repetition beats cramming every time. You're just, what, hoping you remember under pressure? ▷ yields to S2

22.7–28.9 **S2** [turn] But we don't have time for that now! Like, are you seriously starting a whole new study schedule two days before finals? ▷ yields to S1

28.8–35.8 **S1** [turn] I'm not saying start now, I'm just saying your method is pointless long-term. You'll just blank if you get a weird question.

32.1–34.1 **S2** [failed claim] No, because if I know the key words—

35.9–42.2 **S1** [turn] —Key words don't help if it's an application question, like that whole photosynthesis essay last year. ▷ yields to S2

39.3–39.9 **S2** [laughter] haha

42.6–47.9 **S2** [turn] Um, actually, I did fine on that. I just memorized the structure and like, plugged in the details. ▷ yields to S1

48.0–54.4 **S1** [turn] But that's just regurgitating stuff. You can't really explain why, say, C4 plants need that pathway. ▷ cut off

53.3–55.6 **S2** [interruption] Yes I can! It's because they live in, uh— ▷ takes floor from S1

55.7–62.9 **S2** [turn] —they live in hot environments so photorespiration would waste energy if they used normal Calvin cycle, that's why. ▷ yields to S1

62.9–66.5 **S1** [turn] So what's the actual mechanism? Not just the buzzwords. ▷ yields to S2

67.0–68.9 **S2** [turn] You're seriously quizzing me right now? ▷ yields to S1

68.8–74.5 **S1** [turn] Yeah, I am. Because it proves my point. If you just memorize, you can't connect the dots. ▷ holds, pauses

71.0–71.8 **S2** [failed claim] Hold on—

75.3–77.4 **S1** [turn] —No, like, what enzyme starts it? ▷ yields to S2

77.7–85.3 **S2** [turn] PEP carboxylase. See, I do know it. You just have this idea that, I dunno, unless you do it your way it's not real learning or something. ▷ yields to S1

85.9–91.5 **S1** [turn] I mean, it's just the only way that works long-term. No offense, but your approach is risky. ▷ yields to S2

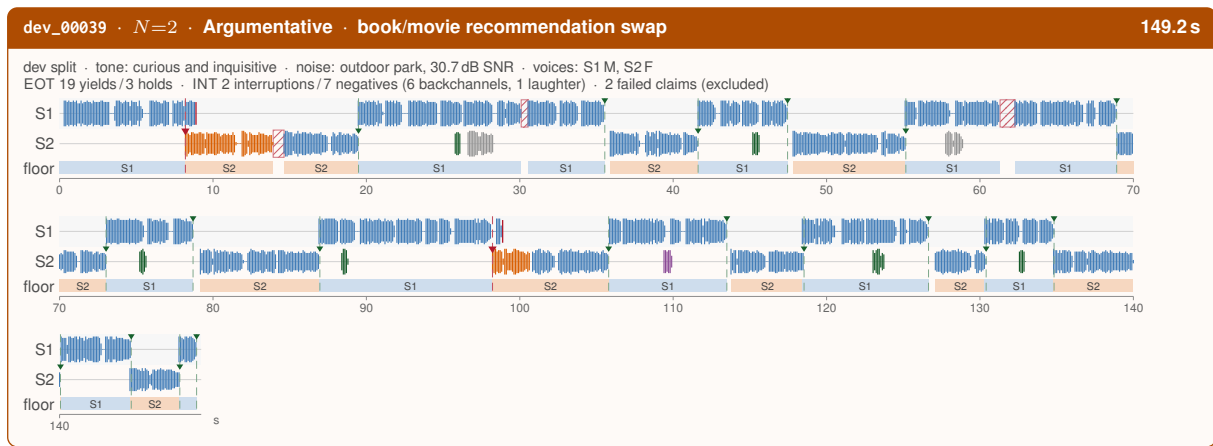
91.9–97.7 **S2** [turn] Right, but I'm not you. Some people need the pressure to focus. Maybe you just like being stressed in advance. ▷ yields to S1

95.7–96.1 **S1** [laughter] heh

97.6–103.1 **S1** [turn] Maybe, but I don't forget everything the second the test's done. You'll see when the grade comes out. ▷ yields to S2

103.1–107.0 **S2** [turn] We'll see. But if I get a higher grade I'm rubbing it in your face, just a warning. ▷ yields, end

105.3–105.8 **S1** [backchannel] mm-hm



Transcript of dev_00039

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.9 **S1** [turn] If you honestly want something that'll stick with you for days, 'Never Let Me Go' is my pick. Ishiguro's style, it's subtle but devastating, and the structure just, I don't know, it lingers. ▷ cut off

8.2–13.9 **S2** [interruption] Wait, but didn't you say before you were tired of dystopian stuff? That one's basically the definition of bleak. ▷ takes floor from S1 ▷ holds, pauses

14.6–19.5 **S2** [turn] Because honestly, I just found it emotionally manipulative, like it's sad for the sake of sadness. ▷ yields to S1

19.5–30.1 **S1** [turn] No, it's not just that, it's—there's complexity. It explores the ethics of, you know, what it means to be human, more than your standard dystopia. It's not 'Hunger Games.' ▷ holds, pauses

25.8–26.2 **S2** [backchannel] mm-hm

26.6–28.3 **S2** [failed claim] I get that, but still, it's like—

30.5–35.5 **S1** [turn] —and it doesn't give easy answers. That's what I'm saying. It's ambiguous to the end. ▷ yields to S2

35.9–41.6 **S2** [turn] But ambiguity doesn't mean it's good, right? Like, don't you get frustrated with books that just, uh, leave you hanging? ▷ yields to S1

41.6–47.5 **S1** [turn] No, that's the point. If it wraps up too neat, it's artificial. Real life isn't like that. ▷ yields to S2

45.2–45.7 **S2** [backchannel] uh-huh

47.8–55.2 **S2** [turn] Well, if we're talking about something that stays with you, I'd say 'Arrival'—the movie, not the short story collection—did that for me way more effectively. ▷ yields to S1

55.1–61.3 **S1** [turn] But the movie, it oversimplifies the linguistics. Villeneuve cuts all the thought experiment stuff out. ▷ holds, pauses

57.8–58.9 **S2** [failed claim] No, it doesn't—

62.3–68.9 **S1** [turn] —in the Ted Chiang story, language literally changes her perception of time, it's, like, way more radical. ▷ yields to S2

68.9–73.0 **S2** [turn] But the emotional arc is better in the movie! That's what matters to most people. ▷ yields to S1

73.1–78.7 **S1** [turn] Well, sure, but if you want your brain challenged, the book smokes the film. It's not even close. ▷ yields to S2

75.2–75.7 **S2** [backchannel] yeah

79.2–87.0 **S2** [turn] Okay but, like, doesn't accessibility count for something? You always pick complicated stuff but I want recommendations I can actually unwind with. ▷ yields to S1

86.9–91.9 **S1** [turn] You say complicated, I say layered. Books that are too easy don't stick with you.

88.4–88.8 **S2** [backchannel] right

92.0–98.9 **S1** [turn] So if you want something more chill, I could name, I don't know, 'Eleanor Oliphant Is Completely Fine', but you'd probably say that one's— ▷ cut off

98.2–100.7 **S2** [interruption] Predictable! Thank you. That's what I mean, you can't— ▷ takes floor from S1

100.8–105.8 **S2** [turn] You can't have it both ways. You can't say you love ambiguity and knock something for being accessible. ▷ yields to S1

105.8–113.5 **S1** [turn] I'm not, though! I'm literally just saying you have to meet a book halfway. Otherwise, you're just scrolling Netflix for the easiest option. ▷ yields to S2

109.4–109.9 **S2** [laughter] haha

113.8–118.5 **S2** [turn] Okay, fine, what about 'Before Sunrise'? If we're swapping, I'd trade your Ishiguro for that. ▷ yields to S1

118.4–126.6 **S1** [turn] That's a trade I'd actually take. But only because Linklater does conversational ambiguity better than almost anyone. ▷ yields to S2

123.0–123.8 **S2** [backchannel] exactly

127.1–130.4 **S2** [turn] See, you do like movies that don't wrap up too neat. Same thing. ▷ yields to S1

130.3–134.8 **S1** [turn] Yeah, but for the record, I still think 'Never Let Me Go' is more haunting. ▷ yields to S2

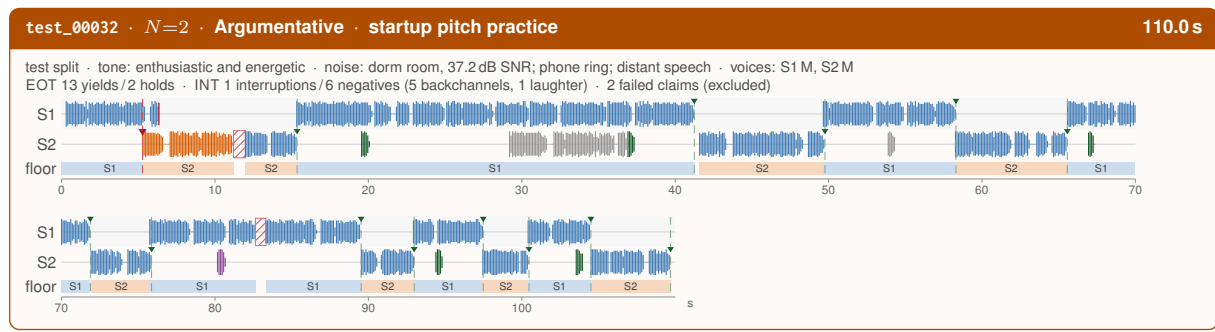
132.6–133.0 **S2** [backchannel] mm-hm

134.8–140.1 **S2** [turn] Fine, we'll call it a draw, but you're reading 'Before Sunrise' as soon as you're done re-brooding over Ishiguro. ▷ yields to S1

140.1–144.7 **S1** [turn] Deal. And I will absolutely text you right after with a ten-point analysis. ▷ yields to S2

144.6–147.8 **S2** [turn] Oh god, I can't wait. Please keep it under a thousand words. ▷ yields to S1

147.8–148.9 **S1** [turn] No promises. ▷ yields, end



Transcript of test_00032

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.3 **S1** [turn] Um, so, basically, our platform matches local chefs with people who want, like, home-cooked meals delivered. It's, uh, kind of like Airbnb meets DoorDash, but, uh, for food? I think there's a huge market for this. ▷ cut off

5.3–11.2 **S2** [interruption] But where's the differentiation? You're saying it's Airbnb meets DoorDash, but why would someone not just use, you know— ▷ takes floor from S1 ▷ holds, pauses

12.0–15.4 **S2** [turn] —DoorDash or Uber Eats? I'm not hearing what's actually new here. ▷ yields to S1

15.3–25.0 **S1** [turn] No, but see, it's different because we're letting completely independent cooks prepare their own menus, not just restaurants. People want, like, authentic food you can't get anywhere else—

19.6–20.1 **S2** [backchannel] uh-huh

25.1–30.4 **S1** [turn] —and it's, I don't know, it's a way to support the community, like, side hustles for people who love to cook.

29.2–36.8 **S2** [failed claim] Okay, but have you actually looked into the legalities? There's a lot of health code stuff involved when you're selling cooked food from homes, it's not—

30.6–35.5 **S1** [turn] —yeah, I know, but we'd have guidelines, kind of like how Airbnb deals with city regulations, right?

35.6–41.2 **S1** [turn] Plus, there are ways to vet chefs, like background checks or maybe a video kitchen tour before they get listed. ▷ yields to S2

36.9–37.4 **S2** [backchannel] Right

41.6–49.8 **S2** [turn] But do you really think people would trust that? I mean, if you see someone's kitchen on video, does that make you comfortable eating their food? I'm not convinced. ▷ yields to S1

49.7–57.0 **S1** [turn] Well, look, people already use home cooks on Instagram all the time! I mean, it's risky, but they do it. If there's a platform with reviews and—

53.9–54.4 **S2** [failed claim] Wait—

57.1–58.3 **S1** [turn] —some guarantee— ▷ yields to S2

58.3–65.6 **S2** [turn] No, hang on, you can't just point to Instagram and say that covers liability—what if someone gets sick? Who gets sued? You or the chef? ▷ yields to S1

65.6–71.9 **S1** [turn] We'd have waivers, I guess? Or insurance built into the fee? I mean, Uber takes on some legal risk, right? It's doable. ▷ yields to S2

67.0–67.4 **S2** [backchannel] mm-hm

71.9–75.9 **S2** [turn] But Uber drivers aren't preparing food. There's a difference, big difference. ▷ yields to S1

75.7–82.6 **S1** [turn] Yeah, yeah, but that's, like, why the opportunity's there. No one else is willing to solve that headache. If we crack that, it could be huge. ▷ holds, pauses

80.2–80.7 **S2** [laughter] haha

83.3–89.5 **S1** [turn] I mean, what if we start with pop-up events instead, build trust, then move to the app full-scale after some pilots? ▷ yields to S2

89.5–93.0 **S2** [turn] That's something, at least. Still think it'll be hard to scale, but maybe. ▷ yields to S1

92.9–97.5 **S1** [turn] You gotta risk a little, right? Maybe we just pilot it on campus so it's more controlled. ▷ yields to S2

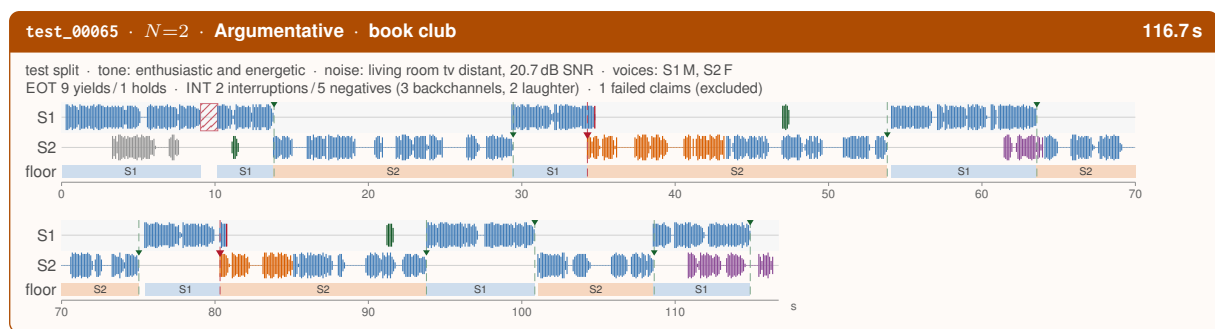
94.4–94.9 **S2** [backchannel] Yeah

97.5–100.5 **S2** [turn] Just, make sure you're not underestimating the red tape, okay? ▷ yields to S1

100.4–104.5 **S1** [turn] I mean, yeah, obviously. But if we don't try, it's never gonna exist, you know? ▷ yields to S2

103.5–104.0 **S2** [backchannel] right

104.5–109.7 **S2** [turn] Look, I'm not saying give up, but you need a real plan for the legal side, or investors won't touch it. ▷ yields, end



Transcript of test_00065

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.1 **S1** [turn] Alright, so, um, I guess I'll just start by saying I thought the pacing of the book was, like, really uneven. The first hundred pages dragged so much, I almost gave up honestly. ▷ holds, pauses

3.3–7.7 **S2** [failed claim] Did you actually read the first hundred pages, or just—

10.2–13.9 **S1** [turn] —Yes, I read them. Very slowly, because it felt like nothing was happening. ▷ yields to S2

11.1–11.6 **S2** [backchannel] mm-hm

13.8–29.4 **S2** [turn] I mean, I get that, but I think you missed all the subtle character work. It's like, you have to actually pay attention to the dialogue. The tension is in the awkward silences. ▷ yields to S1

29.4–34.8 **S1** [turn] But I don't think awkward silences count as tension just because the author refuses to show anyone making a decision. Like, there's a line between subtle and— ▷ cut off

34.3–43.3 **S2** [interruption] —No, see, that’s the point! The indecision is the whole theme. It’s a book about people failing to act. **> takes floor from S1**

43.4–53.8 **S2** [turn] Like, when Ava just sits there at the café and stares at her phone for five pages? That’s modern alienation in a nutshell. **> yields to S1**

47.0–47.5 **S1** [backchannel] uh-huh

54.1–63.6 **S1** [turn] Maybe, but it still made me want to throw the book across the room. I mean, five pages of phone-scrolling? That’s not narrative, that’s... padding. It doesn’t provoke empathy, it just tests my patience. **> yields to S2**

61.4–64.0 **S2** [laughter] Oh wow, patience is a virtue, right?

64.1–75.0 **S2** [turn] But seriously, you’re acting like pacing means constant action. Sometimes you need to, you know, just sit with discomfort. **> yields to S1**

75.4–80.8 **S1** [turn] I’m not saying every book needs car chases, but even the dialogue just went in circles. Like Nate and Ava’s fight about the groceries—how many times do you need to repeat the same argument? **> cut off**

80.3–85.1 **S2** [interruption] Hang on, that’s realistic! Couples do repeat themselves— **> takes floor from S1**

85.1–93.8 **S2** [turn] Like, welcome to actual relationships, you know? It’s not all witty banter and dramatic monologues. **> yields to S1**

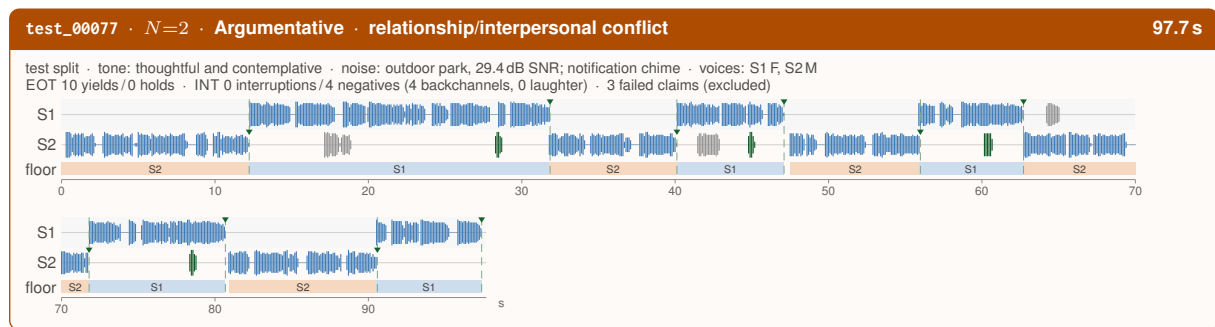
91.2–91.7 **S1** [backchannel] yeah

93.8–100.8 **S1** [turn] But it’s fiction! It doesn’t have to mimic real life to the point of boredom. There’s a responsibility to the reader to, I dunno, edit a little. **> yields to S2**

101.1–108.6 **S2** [turn] But would it even be the same book if you trimmed it down? I think you want a totally different novel. **> yields to S1**

108.6–114.9 **S1** [turn] Maybe I do, honestly. I like stories that actually, you know, move. Or at least don’t make me feel stuck in an endless text thread. **> yields, end**

110.9–116.4 **S2** [laughter] haha, endless text thread is the new realism, get used to it.



Transcript of test_00077

start–end (s) speaker [ground-truth label] text **> outcome**

0.3–12.2 **S2** [turn] Alright, let’s just start from the beginning. I know this is uncomfortable, but I want to make sure both of you get to really explain your side. So, Sarah, can you walk me through what you’re feeling right now? **> yields to S1**

12.3–25.3 **S1** [turn] I mean, honestly, I keep saying the same thing—it feels like I’m the only one making any effort. Like when plans fall through, somehow it’s on me to reschedule, to check in, to keep things afloat. And it’s exhausting.

17.1–18.9 **S2** [failed claim] I get what you’re saying. But do you—

25.4–31.8 **S1** [turn] —I just don’t understand how that’s supposed to be fair. I mean, if the effort isn’t mutual, what’s the point, right? **> yields to S2**

28.3–28.8 **S2** [backchannel] right

31.8–40.1 **S2** [turn] But Sarah, if you always step in first, maybe you’re not giving him the chance to step up, you know? Is it possible that it’s turning into a cycle? **> yields to S1**

40.1–42.9 **S1** [turn] So it’s my fault for actually caring enough to fix things?

41.5–43.0 **S2** [failed claim] No, that’s not what I’m saying—

43.1–47.1 **S1** [turn] —Because if I just sit back and wait, nothing happens. I’ve tried it before. **> yields to S2**

44.8–45.2 **S2** [backchannel] mm-hm

47.5–56.0 **S2** [turn] Can I—okay, I hear you. But how do you know for sure he doesn’t notice or care? Isn’t it possible he just doesn’t express it in the same way? **> yields to S1**

55.9–62.7 **S1** [turn] But if someone cares, they act, right? Like you can’t just say you care and then disappear for three days when things are tense. **> yields to S2**

60.2–60.7 **S2** [backchannel] okay

62.7–65.5 **S2** [turn] But lots of people get really quiet when things are hard—

64.2–65.1 **S1** [failed claim] Sure, but that’s—

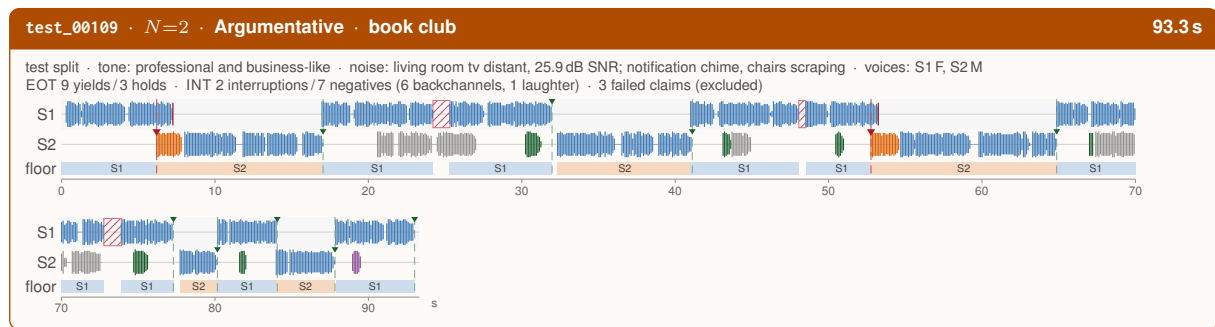
65.7–71.8 **S2** [turn] —not because they don’t care, but because maybe they don’t know what to say. Or they’re worried they’ll make it worse. **> yields to S1**

71.8–80.7 **S1** [turn] But isn’t silence its own statement? I mean, how am I supposed to feel supported if he just vanishes every time things get even a little bit difficult? **> yields to S2**

78.4–78.9 **S2** [backchannel] right

80.9–90.6 **S2** [turn] Would you feel better—like, let’s say next time you just let him reach out, and you wait. No reminders, no extra messages. Or would that feel like giving up for you? **> yields to S1**

90.5–97.4 **S1** [turn] Honestly? I think it would make it painfully clear that I’m the only one who’s been trying. But maybe that’s what I need to see. **> yields, end**



Transcript of test_00109

start–end (s) speaker [ground-truth label] text **> outcome**

0.3–7.3 **S1** [turn] So, I really think “The Quiet City” was trying to make a statement about forgiveness, even if it was a little heavy-handed in how it handled Sarah’s father. It felt to me like the novel wanted us to sympathize, not judge. **> cut off**

6.2–7.9 **S2** [interruption] I'll disagree a bit, actually— **> takes floor from S1**

8.0–17.0 **S2** [turn] I found the way the narrative handled the father almost irresponsible. It gave him redemption with—like, almost no reparation towards Sarah. It's not believable to me. **> yields to S1**

17.0–24.2 **S1** [turn] I mean, but—look, nobody's perfect. The way Sarah struggled with her feelings, that seemed pretty true to life to me. People don't always get closure. **> holds, pauses**

20.6–27.1 **S2** [failed claim] But that's the point, right? The novel skips all the hard conversations. He just shows up at the end and she sort of, just accepts—

25.3–32.0 **S1** [turn] —she accepts him, but I wouldn't call it forgiveness. There's still a lot of distance between them, and I think the author left it that way on purpose. **> yields to S2**

30.2–31.3 **S2** [backchannel] Yeah, maybe.

32.3–41.1 **S2** [turn] But why do you think the mother's storyline basically vanished after chapter seven? It felt like the book just forgot she existed, which made Sarah's family life one-dimensional. **> yields to S1**

41.0–48.1 **S1** [turn] I think that was intentional, though. The absence of the mother becomes another silence Sarah has to deal with, which sort of echoes the main theme. **> holds, pauses**

43.1–43.6 **S2** [backchannel] Hmm.

43.7–45.0 **S2** [failed claim] But the author could've—

48.5–53.2 **S1** [turn] —could've developed it more, sure, but sometimes what's unsaid is just as powerful as what's explicit in the plot. **> cut off**

50.5–51.0 **S2** [backchannel] Right.

52.8–54.6 **S2** [interruption] But I want to push back on that a little— **> takes floor from S1**

54.7–64.9 **S2** [turn] Because, like, for me, the lack of resolution just started to feel like avoidance, not subtlety. I kept waiting for a real confrontation between Sarah and her father and, uh, then it just... fizzled out. **> yields to S1**

64.9–72.8 **S1** [turn] But isn't that how real life works sometimes? I mean, sometimes you go in circles with family and there's no big cathartic moment. That's what I liked, honestly. **> holds, pauses**

67.0–67.3 **S2** [backchannel] mm

67.4–72.6 **S2** [failed claim] But fiction isn't real life, it's an artificial construct. It's supposed to illuminate those moments—

73.9–77.3 **S1** [turn] —right, but maybe it's pointing out that sometimes silence is the message. **> yields to S2**

74.7–75.7 **S2** [backchannel] Hmm... maybe.

77.7–80.2 **S2** [turn] I'm not convinced, but I see where you're coming from. **> yields to S1**

80.2–84.1 **S1** [turn] That's fair. I do wish we'd gotten more from the supporting characters, I'll give you that. **> yields to S2**

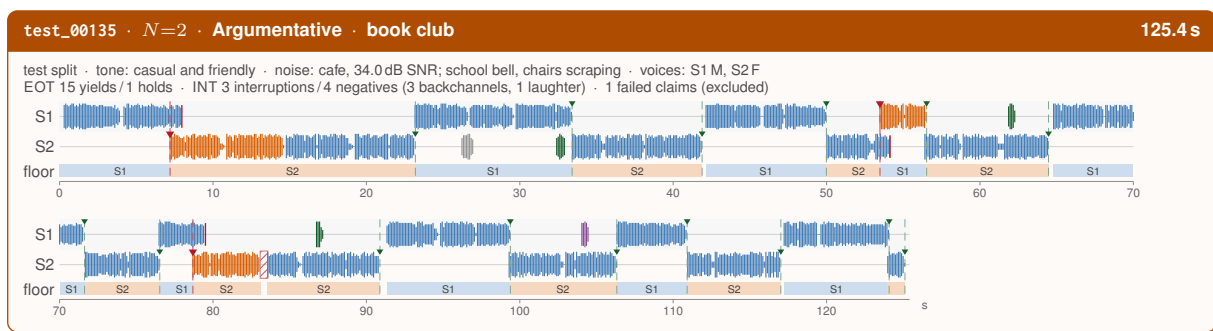
81.6–82.1 **S2** [backchannel] yeah

84.0–87.8 **S2** [turn] And for the record, I thought Sarah's internal monologues were very well-written. **> yields to S1**

87.9–91.0 **S1** [turn] Absolutely! That part was, uh, actually my favorite element.

89.0–89.5 **S2** [laughter] haha

91.2–93.0 **S1** [turn] Well, we can agree on that at least. **> yields, end**



Transcript of test_00135

start-end (s) speaker [ground-truth label] text **> outcome**

0.3–8.0 **S1** [turn] Honestly, I just don't get why everyone loves 'Gone Girl' so much. The characters are all totally unlikable and the twists feel, I don't know, kinda forced to me. **> cut off**

7.2–14.6 **S2** [interruption] Wait, I mean—Nick isn't supposed to be likable, that's the point. Gillian Flynn does it on purpose to make you question your own judgment. **> takes floor from S1**

14.8–23.2 **S2** [turn] You can't say the whole book falls flat just because you don't like the characters. It's about how messed up people can be together, not about rooting for someone. **> yields to S1**

23.2–29.6 **S1** [turn] But if I hate everyone in the book, why would I care what happens to them? It just made me want to put it down every few chapters.

26.2–27.0 **S2** [failed claim] Yeah but—

29.8–33.4 **S1** [turn] —Like, Amy's whole cool girl monologue, it just felt, I don't know, preachy? **> yields to S2**

32.4–33.0 **S2** [backchannel] mm-hm

33.4–41.9 **S2** [turn] But that monologue became iconic! She's calling out all these tropes about women, it's supposed to feel a bit preachy because she's literally laying it bare. **> yields to S1**

42.1–50.0 **S1** [turn] I don't know, maybe I just want someone to root for, even if they're terrible. Like in 'Breaking Bad', Walter's awful but you still kinda want him to win. **> yields to S2**

50.0–54.1 **S2** [turn] But isn't that exactly what makes 'Gone Girl' different? You're not meant to root for anyone, you're just uncomfortable the whole time— **> cut off**

53.5–56.5 **S1** [interruption] Yeah but does that make it good? Like, discomfort for its own sake— **> takes floor from S2 > yields to S2**

56.4–64.5 **S2** [turn] No, no, it's not just to make you squirm! It's to make you think about the lies people live with. The whole unreliable narrator thing was brilliant. **> yields to S1**

61.9–62.3 **S1** [backchannel] uh-huh

64.8–71.6 **S1** [turn] Okay, fine, but did you actually buy Amy's whole plan? Like, that she could frame Nick so perfectly and nobody would catch on? **> yields to S2**

71.7–76.5 **S2** [turn] Well, I think it was extreme, sure, but it's a thriller, it's supposed to push those boundaries. **> yields to S1**

76.5–79.5 **S1** [turn] Sounds kinda like you're making excuses for the plot holes though. **> cut off**

78.7–83.1 **S2** [interruption] Excuse me, no! I just think thrillers don't have to be realistic to be smart. **> takes floor from S1 > holds, pauses**

83.6–90.9 **S2** [turn] It's not a documentary, it's fiction! She's an unreliable narrator, so maybe even her plan isn't as flawless as she thinks. **> yields to S1**

86.8–87.2 **S1** [backchannel] right

91.3–99.4 **S1** [turn] Still, I'd rather have a story where the plot actually holds up. Like, with 'The Girl with the Dragon Tattoo', even the crazy parts fit together better. **> yields to S2**

99.3–106.3 **S2** [turn] But come on, that book was full of wild coincidences too! If you pick at any thriller you'll find stuff that doesn't make sense. **> yields to S1**

104.0–104.5 **S1** [laughter] haha

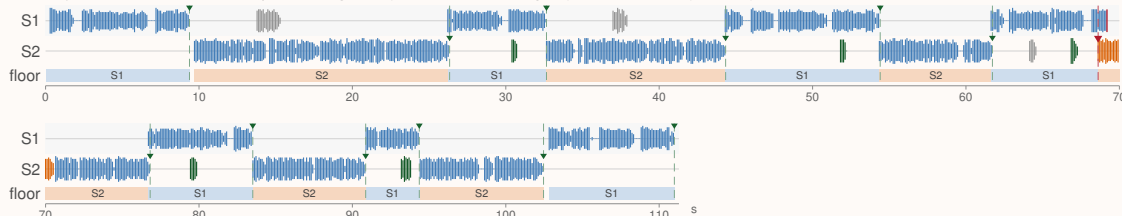
106.3–110.9 **S1** [turn] Yeah, yeah, but at least I didn't want to yeet every single character off a cliff in that one. **> yields to S2**

110.9–117.0 **S2** [turn] Oh my god, you're impossible! Look, even though you hated them, you're still talking about them, right? That says something. ▷ yields to S1
 117.2–124.1 **S1** [turn] Guess that's true. Fine, I'll give you that. But next time can we please pick something with, like, one person who isn't a sociopath? ▷ yields to S2
 124.0–125.1 **S2** [turn] Haha, deal! ▷ yields, end

test_00139 · $N=2$ · Argumentative · classroom discussion

111.3 s

test split · tone: professional and business-like · noise: living room tv distant, 29.5 dB SNR; distant speech · voices: S1 M, S2 M
 EOT 12 yields / 0 holds · INT 1 interruptions / 5 negatives (5 backchannels, 0 laughter) · 3 failed claims (excluded)



Transcript of test_00139

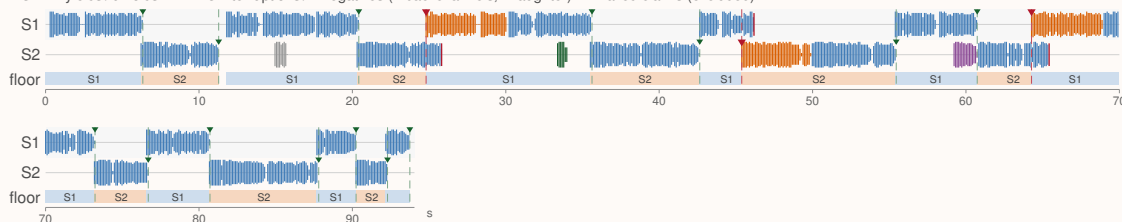
start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.4 **S1** [turn] All right, let's get started. Today's topic is whether mandatory uniforms improve school discipline. I want to hear your honest thoughts, Daniel. ▷ yields to S2
 9.7–17.8 **S2** [turn] Um, well, I think in some cases uniforms could, like, reduce bullying maybe, but I don't think it's a straightforward fix for discipline issues.
 13.8–15.3 **S1** [failed claim] But that's a common argument—
 17.9–26.3 **S2** [turn] —I mean, sorry, but it's just, sometimes the real problems are, uh, like teachers not being consistent or, you know, bigger socioeconomic stuff. ▷ yields to S1
 26.2–32.7 **S1** [turn] Right, but if the uniform policy sets clear expectations, isn't that already helping structure things more? ▷ yields to S2
 30.4–30.8 **S2** [backchannel] mm-hm
 32.7–39.3 **S2** [turn] I just think it's more complicated. Like, uh, some students actually react against strict rules and act out more.
 37.0–38.0 **S1** [failed claim] But you don't think—
 39.4–44.3 **S2** [turn] —and then parents argue about enforcement, so it kind of creates more tension sometimes. ▷ yields to S1
 44.3–54.4 **S1** [turn] Let me finish, Daniel. What I was going to say is, don't you think structure is also a way to level the field, so everyone's judged on behavior, not clothes? ▷ yields to S2
 51.8–52.2 **S2** [backchannel] yeah
 54.4–61.7 **S2** [turn] There's, uh, some value in that, but I'm just not sure it's worth, like, the loss of self-expression. Kids might feel resentful, you know? ▷ yields to S1
 61.6–65.4 **S1** [turn] Isn't resentment a lesser evil compared to daily bullying though?
 64.2–64.6 **S2** [failed claim] Well—
 65.5–69.2 **S1** [turn] —because every time dress codes come up, it's about minimizing those differences that fuel teasing. ▷ cut off
 66.9–67.3 **S2** [backchannel] right
 68.6–70.5 **S2** [interruption] But I, um, I don't think uniforms alone— ▷ takes floor from S1
 70.7–76.8 **S2** [turn] —actually solve bullying. There's always something else kids will pick on. And sometimes it just goes underground. ▷ yields to S1
 76.7–83.5 **S1** [turn] True, but we have data showing disciplinary referrals drop in districts with uniforms. Doesn't that mean something? ▷ yields to S2
 79.4–79.9 **S2** [backchannel] maybe
 83.5–90.9 **S2** [turn] But sometimes, uh, those stats don't show if students are just, like, holding it in or if teachers are enforcing things differently. ▷ yields to S1
 90.9–94.4 **S1** [turn] So your point is the numbers might not reflect the atmosphere? ▷ yields to S2
 93.2–93.9 **S2** [backchannel] exactly
 94.4–102.5 **S2** [turn] I mean, I guess, if uniforms help some kids feel safer, that's good. But I just, uh, worry it's not, you know, addressing root causes. ▷ yields to S1
 102.8–111.0 **S1** [turn] I still think, practically, it's a positive step. But I appreciate your nuance there, Daniel. Let's open it up to the rest of the class. ▷ yields, end

test_00149 · $N=2$ · Argumentative · restaurant/group dining decision

94.0 s

test split · tone: enthusiastic and energetic · noise: quiet office, 27.4 dB SNR · voices: S1 F, S2 M
 EOT 14 yields / 0 holds · INT 3 interruptions / 2 negatives (1 backchannels, 1 laughter) · 1 failed claims (excluded)



Transcript of test_00149

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.3 **S1** [turn] Alright, so can we just decide on a place? I'm honestly starving and I really don't want to spend another half hour arguing. ▷ yields to S2
 6.2–11.3 **S2** [turn] It's not arguing, it's just—look, pizza's the easiest. You can't tell me you're that against it. ▷ yields to S1
 11.8–17.6 **S1** [turn] I honestly can't do pizza again, Ben. We literally had that like two nights ago, and you know I'm trying to eat less cheese—
 15.0–15.8 **S2** [failed claim] C'mon, it's not—
 17.7–20.4 **S1** [turn] —plus, you never want sushi and I always compromise. ▷ yields to S2
 20.3–25.8 **S2** [turn] But sushi's just so expensive though, and it never fills me up. Like, it's just rice and tiny fish slices— ▷ cut off
 24.8–30.1 **S1** [interruption] That's so not true! You just order two rolls and then complain you're hungry. There's ramen too, you always forget— ▷ takes floor from S2
 30.2–35.6 **S1** [turn] —they have like, huge bowls. You never actually try to pick from the rest of the menu, you just say you're still hungry. ▷ yields to S2
 33.4–34.0 **S2** [backchannel] Right, right.
 35.5–42.6 **S2** [turn] But it's not just that. Pizza's easy to split, no drama. And honestly, I just wanna chill guys, not plan a whole expedition. ▷ yields to S1
 42.6–46.2 **S1** [turn] You're being so dramatic. We literally have three sushi places right near here— ▷ cut off
 45.4–49.9 **S2** [interruption] Hold on, but weren't you the one who got totally grossed out by that last place? Remember? ▷ takes floor from S1
 50.0–55.5 **S2** [turn] You literally said you'd never eat there again because of the wasabi incident. So that only leaves two. ▷ yields to S1

55.4–60.7 **S1** [turn] Haha, okay, that was one bad night, but I'll survive. If we get sushi, I'll even pick up your order, Ben. ▷ yields to S2

59.2–60.8 **S2** [laughter] Haha, wow, bribe me more, I dare you.

60.8–65.4 **S2** [turn] But what if we compromise—hear me out—Indian food! They deliver fast, we haven't done it in months, everyone wins. ▷ cut off

64.3–68.9 **S1** [interruption] Wait, you just said you don't want drama and now you want to negotiate tandoori spice levels all night? ▷ takes floor from S2

69.0–73.2 **S1** [turn] You literally spend a half hour insisting you're okay with heat and then die at medium. ▷ yields to S2

73.2–76.7 **S2** [turn] Man, I'm not that bad. You act like I've never eaten a chili in my life. ▷ yields to S1

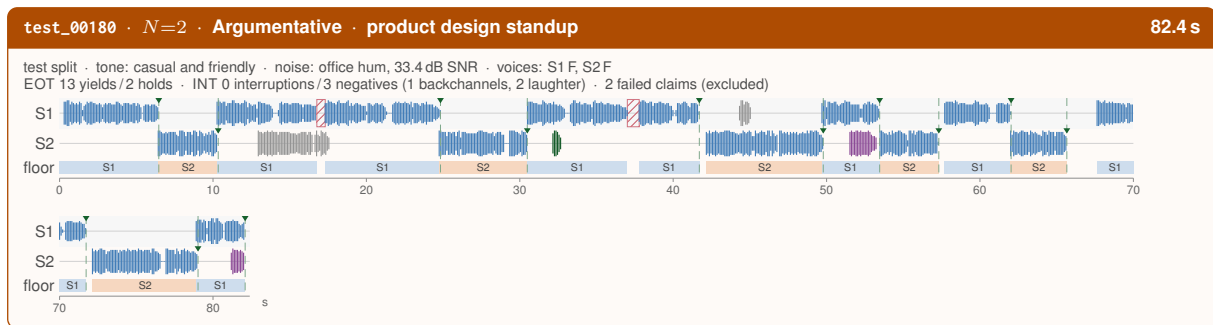
76.6–80.7 **S1** [turn] Oh come on, last time you drank like three glasses of milk and still sulked for an hour. ▷ yields to S2

80.7–87.8 **S2** [turn] Alright, alright. What if, radical idea, we order a little of everything? Pizza, sushi, samosas, just go wild. Feast night. ▷ yields to S1

87.7–90.2 **S1** [turn] Honestly... fine, but you're paying the delivery fees. ▷ yields to S2

90.2–92.3 **S2** [turn] Deal. Anything to get food here before midnight. ▷ yields to S1

92.2–93.7 **S1** [turn] Deal. I'll text the orders. ▷ yields, end



Transcript of test_00180

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.5 **S1** [turn] All right, let's just start—sprint demo's tomorrow, so we need to decide on the onboarding flow today, no more putting it off. ▷ yields to S2

6.4–10.3 **S2** [turn] Why are we changing it again? The last version tested fine with users, right? ▷ yields to S1

10.3–16.8 **S1** [turn] Barely, though. People kept missing the swipe hint, it's buried in that modal. I want to move it to the first screen, way more visible. ▷ holds, pauses

13.0–17.6 **S2** [failed claim] But if you stick it upfront, it's gonna overwhelm users with too much info. That's classic—

17.3–24.8 **S1** [turn] —I'm literally just adding one graphic and two words, not a tutorial pop-up. We have metrics showing half the testers got stuck last round. ▷ yields to S2

24.7–30.5 **S2** [turn] Yeah but the people who got stuck were older users, and they all said tooltips annoyed them. Like, I'm not making that up. ▷ yields to S1

30.5–37.0 **S1** [turn] Older users are not the only segment, though. If we optimize for one group, everyone else just gets a worse experience. ▷ holds, pauses

32.2–32.7 **S2** [backchannel] Right

37.8–41.7 **S1** [turn] So why don't we test the more obvious first? If it's too much, roll back. ▷ yields to S2

42.2–46.9 **S2** [turn] You say 'roll back' but you know that never happens, we always end up stuck with whatever gets shipped first.

44.3–45.1 **S1** [failed claim] That's not—

46.9–49.8 **S2** [turn] No, seriously, name one feature we yanked after launch. ▷ yields to S1

49.7–53.5 **S1** [turn] —the dark mode toggle. We did that last March. But you're right, it's rare. ▷ yields to S2

51.5–53.3 **S2** [laughter] Haha, yeah, that hardly counts.

53.5–57.3 **S2** [turn] Can we please just A/B it quietly? Like, ten percent of new users only. ▷ yields to S1

57.7–62.0 **S1** [turn] Sure, but then don't complain when we need extra QA. It's more overhead. ▷ yields to S2

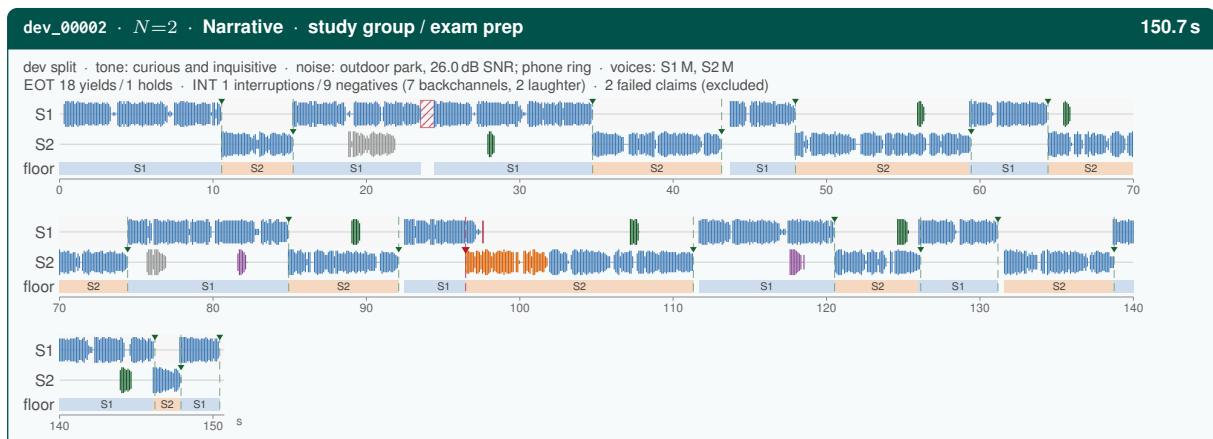
62.0–65.7 **S2** [turn] It's worth it. Remember the search bar change? Saved us so much support time. ▷ yields to S1

67.6–71.7 **S1** [turn] Fine, but I still think the modal isn't clear enough, even for the control group. ▷ yields to S2

72.1–79.0 **S2** [turn] Okay, how about this: we tweak the modal copy AND try your graphic, but only for new installs. That way if it sucks, nobody notices. ▷ yields to S1

78.9–82.1 **S1** [turn] Deal. But if feedback's negative, you own the bug tickets. ▷ yields, end

81.2–82.1 **S2** [laughter] Haha, deal.



Transcript of dev_00002

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–10.6 **S1** [turn] So, remember last semester when I totally blanked on the chem exam? Like, I sat down and instantly forgot everything about titrations. I could see the diagram in my head but the formulas just—poof—gone. ▷ yields to S2

10.6–15.2 **S2** [turn] Yeah, I remember you texting that morning, like 'SOS, what even is an equivalence point?' ▷ yields to S1

15.2–23.5 **S1** [turn] Exactly! And then, I started panicking, which made it even harder to recall. But the weird thing is, after the test, everything came flooding back while I was eating lunch. ▷ holds, pauses

18.9–21.9 **S2** [failed claim] That always happens. It's like your brain trolls you—

24.4–34.8 **S1** [turn] —Right! It's like, why couldn't I remember when it actually mattered? I've started making up little stories or images for the steps. For titration, I picture, uh, a waiter pouring sauce and hitting the perfect spot. ▷ yields to S2

27.9–28.4 **S2** [backchannel] mm-hm

34.7–43.2 **S2** [turn] I do that, too, but I get competitive with it. Like, my trick is, if I get stuck, I just challenge myself to write down everything I know, no matter how random. ▷ yields to S1

43.7–48.0 **S1** [turn] Oh wow, that's smart. So you kinda prime your brain with any fact, and then see what connects? ▷ yields to S2

48.0–59.4 **S2** [turn] Yeah, exactly. One time during finals, I forgot the definition of entropy. I just started jotting, like, 'disorder, thermodynamics, universe gets messier'—and eventually, the full definition popped into my head. ▷ yields to S1

55.9–56.5 **S1** [backchannel] right

59.3–64.4 **S1** [turn] That's actually really helpful. I feel like sometimes I freeze because I'm expecting the answer to just— ▷ yields to S2

64.5–68.4 **S2** [turn] Just show up in your head, yeah. But it doesn't. You have to nudge your brain a bit.

65.5–65.9 **S1** [backchannel] yeah

68.5–74.4 **S2** [turn] So, what about you? Any other weird memory hacks? I remember you said you had a song for the cell cycle or something? ▷ yields to S1

74.5–76.1 **S1** [turn] Yeah, okay, so don't laugh but—

75.7–77.0 **S2** [failed claim] No promises.

76.2–84.9 **S1** [turn] —I literally sing the cell phases to the tune of, like, 'Twinkle Twinkle Little Star.' I go like, 'Interphase, prophase, metaphase and more.' And it just sticks, embarrassingly well. ▷ yields to S2

81.6–82.2 **S2** [laughter] haha

84.9–92.1 **S2** [turn] Honestly, whatever works! I memorized the periodic table by writing it in marker on my bathroom mirror until my mom yelled at me. ▷ yields to S1

89.1–89.7 **S1** [backchannel] no way

92.5–97.6 **S1** [turn] There's something about physical movement, or like, doing something silly, that makes it easier to remember. It's like the goofier, the better. ▷ cut off

96.5–101.8 **S2** [interruption] Totally. Okay, but—have you ever completely bombed despite all that prep? Like, just crashed and burned? ▷ takes floor from S1

102.0–111.3 **S2** [turn] Because I have, for sure. I'll never forget my stats midterm sophomore year. I walked in feeling great, and halfway through I realized I'd memorized the formulas for the wrong test. ▷ yields to S1

107.2–107.8 **S1** [backchannel] oh man

111.7–120.5 **S1** [turn] Oof, that hurts. But yeah, physics last year. I thought I nailed the circuits section but turns out I reversed everything. The TA's note literally said, 'Interesting approach, but not quite.' ▷ yields to S2

117.6–118.6 **S2** [laughter] haha right

120.6–126.1 **S2** [turn] But the wild thing is, after bombing, you remember it forever. Like, I'll never mix up Z-scores again. ▷ yields to S1

124.6–125.4 **S1** [backchannel] exactly

126.0–131.2 **S1** [turn] Honestly, maybe failing is its own study tool? You mess up so badly you have to get it right next time. ▷ yields to S2

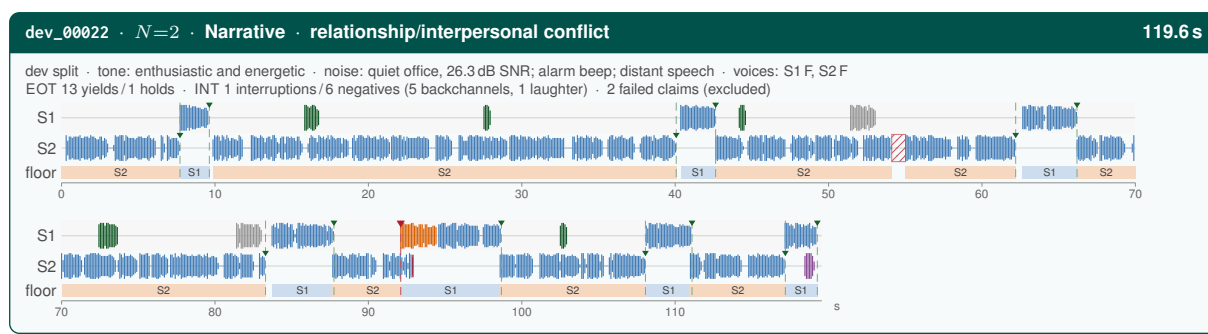
131.6–138.8 **S2** [turn] Yeah, as long as you get another shot, it works. So—uh, what's the weirdest way you've tried to memorize something? Like, stranger than the songs? ▷ yields to S1

138.7–146.2 **S1** [turn] I once drew the entire digestive system on my arm in Sharpie. But then I forgot it was there and went to dinner. The waiter looked so confused. ▷ yields to S2

144.0–144.7 **S2** [backchannel] no way

146.1–147.9 **S2** [turn] That's dedication, dude. ▷ yields to S1

147.8–150.4 **S1** [turn] haha well, whatever gets us through finals week, right? ▷ yields, end



Transcript of dev_00022

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–7.7 **S2** [turn] Okay, so you know how I was telling you about Alex last week? Something wild happened with her and Jamie on Saturday—like, actually wild. ▷ yields to S1

7.7–9.6 **S1** [turn] Oh no, what did they do this time? ▷ yields to S2

9.9–21.4 **S2** [turn] So we're all at Jen's housewarming, right? Everyone's chill, music's loud, people are playing Mario Kart. Suddenly, Jamie just walks out of the kitchen fuming. Like, I honestly thought he was going to throw something.

15.9–16.8 **S1** [backchannel] Wait, really?

21.5–29.1 **S2** [turn] Yeah! He grabs his shoes and just slams the door. Everyone went silent. Alex—she tries to laugh it off, but you could tell she was shaking.

27.5–28.0 **S1** [backchannel] Mm-hm.

29.2–40.1 **S2** [turn] So I followed Jamie outside 'cause honestly I've never seen him like that. He's pacing on the sidewalk, muttering. I ask him what happened and he's like, 'She called me controlling, in front of everyone.' ▷ yields to S1

40.4–42.6 **S1** [turn] Oh wow, she did that in front of everybody? ▷ yields to S2

42.6–49.1 **S2** [turn] Yeah! And it wasn't even, like, a joke. You know when someone's fake-laughing but actually mad? Alex totally did that thing.

44.2–44.6 **S1** [backchannel] Right.

49.2–54.1 **S2** [turn] I mean, who does that? Like, at someone else's party? I tried to get Jamie to go back in but— ▷ holds, pauses

51.4–53.1 **S1** [failed claim] Was he crying or something?

55.0–62.2 **S2** [turn] —not like crying, but you know, red-eyed and super tense. He just walked off. I waited for a while but then I had to go back inside. ▷ yields to S1

62.7–66.2 **S1** [turn] Jeez. That's, uh, pretty rough. Did Alex seem bothered or just pretending? ▷ yields to S2

66.2–75.0 **S2** [turn] She came over to me after, acting like nothing even happened! But when people started leaving, she just went out on the balcony with her phone, not talking to anyone.

72.4–73.7 **S1** [backchannel] Yeah, classic Alex.

75.1–83.3 **S2** [turn] Haha, right? I didn't know what to do. I texted Jamie later, he didn't respond 'til the next morning. All he said was, 'Thanks for coming out.' That's it. ▷ yields to S1

81.4–83.1 **S1** [failed claim] Oof, that's not a great sign.

83.7–87.8 **S1** [turn] Do you think they're going to break up? Or is this, like, just another weird blip for them? ▷ yields to S2

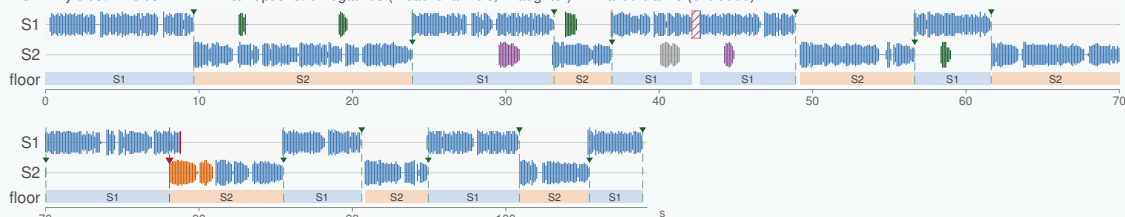
87.7–92.9 **S2** [turn] I honestly don't know. They always fight, but this felt different. I mean, he actually left. Like, gone-gone. ▷ cut off

92.1–94.5 **S1** [interruption] But didn't he do that at Sarah's BBQ too— **▷ takes floor from S2**
 94.6–98.7 **S1** [turn] —remember? He stormed off but came back after an hour. So maybe it's their thing. **▷ yields to S2**
 98.6–103.8 **S2** [turn] Yeah but this time, he didn't come back at all. I dunno, I just feel bad for both of them, honestly.
 102.5–103.0 **S1** [backchannel] Yeah.
 103.9–108.1 **S2** [turn] Like, I wish I could help but... I feel like if I pick a side I'll make it worse. **▷ yields to S1**
 108.1–111.7 **S1** [turn] You probably would. I mean, that's a minefield. I wouldn't touch it. **▷ yields to S2**
 111.0–117.2 **S2** [turn] Haha, yeah. Can I just go back to pretending their drama is, like, a TV show and I'm just an audience member? **▷ yields to S1**
 117.2–119.3 **S1** [turn] Honestly, same. Just get your popcorn. **▷ yields, end**
 118.5–119.1 **S2** [laughter] Haha!

test_00005 · $N=2$ · Narrative · group trivia night

109.2 s

test split · tone: curious and inquisitive · noise: cafe, 34.3 dB SNR; door knock, notification chime; distant speech · voices: S1 F, S2 F
 EOT 14 yields / 1 holds · INT 1 interruptions / 6 negatives (4 backchannels, 2 laughter) · 1 failed claims (excluded)



Transcript of test_00005

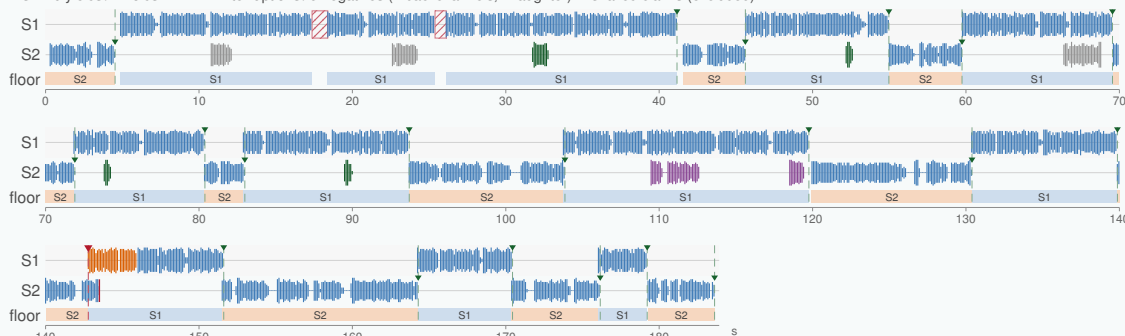
start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.7 **S1** [turn] So, this is our third trivia night—are you feeling lucky this time? I mean last week was rough, I can't believe we missed that question about Pluto. I still haven't forgiven myself for that one. **▷ yields to S2**
 9.7–16.5 **S2** [turn] Yeah, I'm ready, but that Pluto question was sneaky. Technically a dwarf planet. They always throw in those trick definitions, you know?
 12.6–13.1 **S1** [backchannel] Mhm
 16.6–23.9 **S2** [turn] But I did some reading on weird planet facts this week. Just in case. Did you know Venus rotates the opposite direction of most planets? **▷ yields to S1**
 19.1–19.7 **S1** [backchannel] Right
 23.9–33.2 **S1** [turn] Actually, yeah, I learned that from one of my niece's Zoom classes last year. She's obsessed with planets. She made me guess which planet spins backwards for like fifteen solid minutes. **▷ yields to S2**
 29.6–31.0 **S2** [laughter] Haha, relentless.
 33.0–36.9 **S2** [turn] But that's exactly the kind of random detail these trivia nights love. **▷ yields to S1**
 33.9–34.7 **S1** [backchannel] Exactly
 36.9–42.1 **S1** [turn] Actually, the one that always gets me is the sports round. I just blank out. Like, who scored the most points in— **▷ holds, pauses**
 40.1–41.4 **S2** [failed claim] —LeBron, right?
 42.7–48.9 **S1** [turn] —NBA history, yeah, but somehow I always write Michael Jordan out of habit! My dad would have my head for that, honestly. **▷ yields to S2**
 44.3–44.9 **S2** [laughter] Haha
 49.2–56.7 **S2** [turn] You know, I got publicly shamed at trivia a few years ago for forgetting which city hosts the Kentucky Derby. Like, of all the things to miss. **▷ yields to S1**
 56.6–61.7 **S1** [turn] Oof, Louisville! I always remember because my aunt sends me these ridiculous horse-themed coffee mugs. **▷ yields to S2**
 58.4–59.0 **S2** [backchannel] Right
 61.7–70.0 **S2** [turn] That's a pretty memorable way to learn. Honestly, maybe we should each just pick one weird topic to become experts on. Like you do animals, I'll do sports. **▷ yields to S1**
 70.0–78.8 **S1** [turn] Deal. Except, you know, I panic whenever the animal questions get too technical. Last month—uh, the question about what mammal lays eggs? I totally forgot about the echidna and just wrote platypus. **▷ cut off**
 78.1–81.0 **S2** [interruption] Me too! That always gets people. Wait, actually— **▷ takes floor from S1**
 81.1–85.5 **S2** [turn] —I think I read there's a third monotreme? Some weird Australian thing, maybe? **▷ yields to S1**
 85.5–90.6 **S1** [turn] Uh, is there? I thought it was just echidna and platypus. Unless, are there multiple types of echidna? **▷ yields to S2**
 90.8–95.0 **S2** [turn] Yeah, maybe like several species of echidna. I should look it up before tonight. **▷ yields to S1**
 94.9–100.9 **S1** [turn] Smart. Okay, but if there's a music round, I'm counting on you. My brain just shuts off when anyone says 'classic rock'. **▷ yields to S2**
 100.9–105.5 **S2** [turn] Haha, no pressure. But if anything about 70s bands comes up, I'm your girl. **▷ yields to S1**
 105.4–108.9 **S1** [turn] Perfect. Okay, bring on the weird questions—I'm ready this time. Probably. **▷ yields, end**

test_00014 · $N=2$ · Narrative · tech support call

183.9 s

test split · tone: casual and friendly · noise: office hvac, 29.1 dB SNR; school bell · voices: S1 F, S2 M
 EOT 20 yields / 2 holds · INT 1 interruptions / 6 negatives (4 backchannels, 2 laughter) · 3 failed claims (excluded)



Transcript of test_00014

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–4.5 **S2** [turn] Thank you for calling tech support, this is Alex. How can I help you today? ▷ yields to S1

4.9–17.4 **S1** [turn] Hey Alex, thanks for answering so quick. So, um, I have this laptop that's been acting up all day—like, it just randomly freezes whenever I try to open my email or, honestly, sometimes even when I'm just scrolling. ▷ holds, pauses

10.8–12.2 **S2** [failed claim] Oh no, that's frustrating.

18.4–25.4 **S1** [turn] Yeah, it's, um, kind of ruining my workflow. Like, I had this big document I was updating for work, and I swear, I lost— ▷ holds, pauses

22.6–24.3 **S2** [failed claim] Did it crash in the middle of saving?

26.1–34.4 **S1** [turn] —like, half an hour's work right there. Yeah, exactly! It just froze and then when it restarted, the auto-save hadn't even kicked in for some reason.

31.8–32.8 **S2** [backchannel] Yikes. That's rough.

34.5–41.2 **S1** [turn] Right? It's happened before, but today's been just—like, non-stop. I'm honestly scared to open anything, even my calendar. ▷ yields to S2

41.6–45.6 **S2** [turn] So, just to get a picture, is this a work laptop or your personal one? ▷ yields to S1

45.6–55.0 **S1** [turn] It's my work laptop, yeah. I've only had it for maybe a year? IT set it up originally, but I haven't had major problems until, like, last week. And today's been the worst. ▷ yields to S2

52.2–52.7 **S2** [backchannel] Mm-hm.

55.0–59.7 **S2** [turn] Has anything changed recently—like, new updates, new software you installed? ▷ yields to S1

59.8–69.6 **S1** [turn] Oh actually, now that you mention it, Windows did one of those forced updates a few nights ago. I think it was, um, Tuesday? Afterwards, my browser started acting funny too. ▷ yields to S2

66.4–68.9 **S2** [failed claim] Yeah, sometimes those updates mess with stuff.

69.6–71.9 **S2** [turn] Has it only been Chrome, or other apps too? ▷ yields to S1

71.9–80.4 **S1** [turn] It's definitely Chrome that freezes most, but I noticed Outlook did it once. And, oh, even the file explorer, when I tried to move some PDFs— ▷ yields to S2

73.8–74.3 **S2** [backchannel] Yeah.

80.4–83.0 **S2** [turn] Were you plugged in or on battery when that happened? ▷ yields to S1

82.9–93.7 **S1** [turn] Plugged in, actually. It's almost always plugged in because the battery is, like, not great anymore. But I used to be able to go to the coffee shop and work there, y'know, but now I just don't trust it. ▷ yields to S2

89.5–90.0 **S2** [backchannel] Right.

93.7–103.9 **S2** [turn] You know, that kind of sudden freezing usually makes me think of either a memory issue or sometimes overheating. Do you notice the fans getting loud or it feeling hot? ▷ yields to S1

103.8–113.7 **S1** [turn] Oh, totally. Sometimes it sounds like a jet engine. It definitely gets warm on my lap. I brushed the vents out with, like, a Q-tip one time, but I'm not sure it made a difference.

109.5–112.7 **S2** [laughter] Haha, yeah, those laptop fans can make you paranoid.

113.9–119.8 **S1** [turn] Seriously! My cat actually hides when she hears it start up. She thinks something's about to explode. ▷ yields to S2

118.5–119.5 **S2** [laughter] Haha, that's great.

120.0–130.4 **S2** [turn] Okay, so what I'd recommend first is checking task manager to see what's using up all the resources. Um, can you open that for me and tell me what's at the top? ▷ yields to S1

130.4–139.9 **S1** [turn] Sure, hang on—okay, let me—oh, yeah, Chrome is, uh, using, like, half the memory, and there's something called "Antimalware Service Executable" taking a bunch of CPU? ▷ yields to S2

139.9–143.5 **S2** [turn] Yeah, that's just Windows Defender. It can get a little aggressive after updates. ▷ cut off

142.8–146.0 **S1** [interruption] So, wait, do I just close Defender? Or what should I do? ▷ takes floor from S2

146.0–151.6 **S1** [turn] I mean, I don't wanna, like, break security or anything, but if it's making everything else freeze— ▷ yields to S2

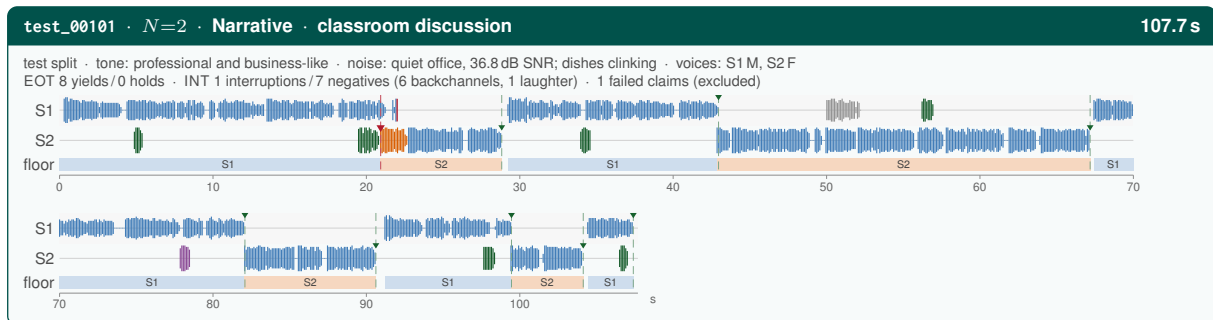
151.5–164.3 **S2** [turn] No, no, definitely don't shut it off entirely. What we can try is a quick restart, and then we can run a disk cleanup. Sometimes that clears out whatever's clogging things up after a big update. ▷ yields to S1

164.2–170.4 **S1** [turn] Okay, I can try that. Honestly, if this fixes it, you'll be my hero. I've been stressing about this all day. ▷ yields to S2

170.4–176.2 **S2** [turn] Let's give it a shot. And if it gives you any trouble, just call back—I'll even put a note in your ticket. ▷ yields to S1

176.0–179.2 **S1** [turn] Thank you, Alex, seriously. I'm gonna cross my fingers. ▷ yields to S2

179.3–183.6 **S2** [turn] No problem. Good luck—and, uh, I hope your cat survives the fan noise! ▷ yields, end



Transcript of test_00101

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–13.4 **S1** [turn] Alright, so—uh, I wanted to start by sharing a moment from my student teaching days. I remember this one class where everything just completely fell apart, but in like, the best possible way. You know? It was supposed to be just a regular lesson on ecosystems, but—

4.9–5.5 **S2** [backchannel] mm-hm

13.5–22.0 **S1** [turn] —uh, one student just blurted out, 'But how do we know the rainforest even matters to us here?' And suddenly—I mean, almost the whole class starts debating. At first I was worried the discussion wouldn't go anywhere, but honestly, it was the liveliest they'd ever been. ▷ cut off

19.5–20.9 **S2** [backchannel] That's fantastic.

20.9–22.7 **S2** [interruption] Did you let them just run with it? ▷ takes floor from S1

22.8–28.8 **S2** [turn] Because honestly, those are the moments when you see them making connections, right? Even if it means skipping your planned activities. ▷ yields to S1

29.2–43.0 **S1** [turn] Exactly! I totally scrapped the worksheet and just guided them with questions. It was so much more engaging. One kid even brought up air quality—he'd heard something on the news—and suddenly we're talking about local pollution too. It made me rethink my own approach to lesson planning. ▷ yields to S2

34.0–34.6 **S2** [backchannel] Yeah

42.9–52.0 **S2** [turn] I know what you mean—I had a section last year in sixth grade where I tried to stick to the book for a unit on government, and it just... fizzled. But as soon as I brought up local elections—

50.0–52.2 **S1** [failed claim] Uh, sorry, so they actually cared about that?

52.1–57.7 **S2** [turn] —yeah, they really did! Because, you know, their parents were voting and it was on their minds. Suddenly everyone had an opinion.

56.2–57.0 **S1** [backchannel] Right, right.

57.8–67.2 **S2** [turn] And it turned into this great, kind of spontaneous civic engagement project. I honestly learned as much as they did. Sometimes I think those moments are more valuable than any test prep. ▷ yields to S1

67.4–82.1 **S1** [turn] For sure, I mean—haha—sometimes I wish administrators could see how powerful it is when you just let a conversation breathe. I had a supervisor drop in once right in the middle of one of those discussions, and I thought I was in trouble, but she actually praised how involved everyone was. ▷ yields to S2

77.9–78.5 **S2** [laughter] haha

82.1–90.6 **S2** [turn] It's so true, and I think that's something newer teachers worry about a lot—"Am I losing control of the class?" But really it's just a different kind of structure, if you trust the process. [▷ yields to S1](#)

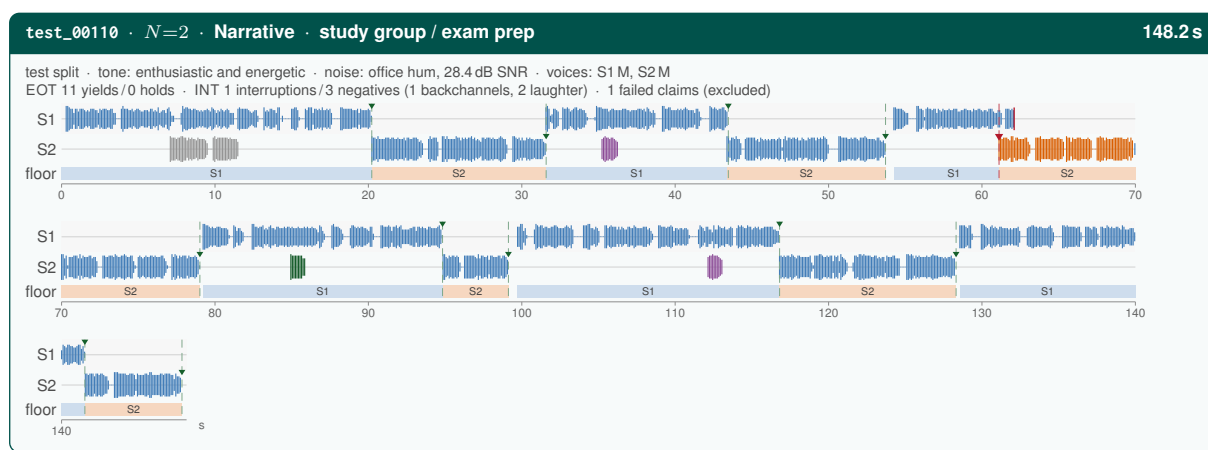
91.2–99.5 **S1** [turn] Yeah, totally, and the students remember those days. I had one reach out last year—he said he still thought about the "rainforest argument." That just made my week. [▷ yields to S2](#)

97.7–98.4 **S2** [backchannel] Oh wow

99.4–104.1 **S2** [turn] That's the stuff that sticks, honestly. I think those are the moments that make me want to keep teaching. [▷ yields to S1](#)

104.4–107.4 **S1** [turn] Same here. It reminds me why all the tough days are worth it. [▷ yields, end](#)

106.5–107.1 **S2** [backchannel] mm-hm



Transcript of test_00110

start–end (s) speaker [ground-truth label] text [▷ outcome](#)

0.3–20.2 **S1** [turn] Alright, before we dive into the flashcards, I gotta tell you, last semester's bio exam totally caught me off guard. I was so sure there'd be a massive essay on mitochondria, so I memorized, like, every random fact. And then... it was just one short answer. I don't even know why I did that to myself. [▷ yields to S2](#)

7.1–11.6 **S2** [failed claim] Oh wow, yeah, that kind of thing always throws me off. Like, you prepare for one thing and then—

20.2–31.6 **S2** [turn] I uh, I usually end up spending way too much time on the diagrams. Last time? I spent at least two hours drawing out a perfect nephron, and then all they wanted was the function. I could've just written filtration and moved on. [▷ yields to S1](#)

31.6–43.5 **S1** [turn] Right? Dude, I did that for the heart diagram. I labeled every single little vein and then there was just this question, like, What's the difference between an artery and a vein? I was so salty. [▷ yields to S2](#)

35.2–36.3 **S2** [laughter] Haha, yeah.

43.4–53.7 **S2** [turn] That's, like, classic. You know, my high school chemistry teacher used to say, 'Don't study what you want to study, study what you're afraid they'll ask.' But I never learn, I always try to master the weirdest topics. [▷ yields to S1](#)

54.3–62.1 **S1** [turn] Oh absolutely. I mean, it's way more fun memorizing all the enzymes in glycolysis than actually, like, remembering their main function. But hey, did you ever bomb a test so hard it just turned funny? Because— [▷ cut off](#)

61.1–69.9 **S2** [interruption] Yeah—wait, I mean, actually I totally have. It was sophomore year, physics midterm. I completely blanked on Ohm's Law. Like, I wrote V equals IR backwards in my head? [▷ takes floor from S1](#)

70.0–79.0 **S2** [turn] And I kept doing all the calculations that way! Only realized after, when my friend was like, 'Dude, why is your resistance negative?' I laughed so hard I started crying. [▷ yields to S1](#)

79.2–94.8 **S1** [turn] Haha, man, that's amazing. I feel you. Uh, my worst was probably, um, when I mixed up meiosis and mitosis on the final. Whole essay. Realized halfway through. Just committed, finished it out, and drew the worst chromosome diagram ever. [▷ yields to S2](#)

84.9–85.9 **S2** [backchannel] Oh no way.

94.9–99.1 **S2** [turn] That's, uh... That's rough. But yeah, I guess, after something like that, you just have to laugh. [▷ yields to S1](#)

99.7–116.8 **S1** [turn] Exactly. Like, at some point, it's just out of your control, right? But, um, okay, have you ever had a teacher who just made everything sound, like, way more dramatic than it needed to be? Mrs. Wu, our calculus teacher, would shout derivatives like she was in an action movie. [▷ yields to S2](#)

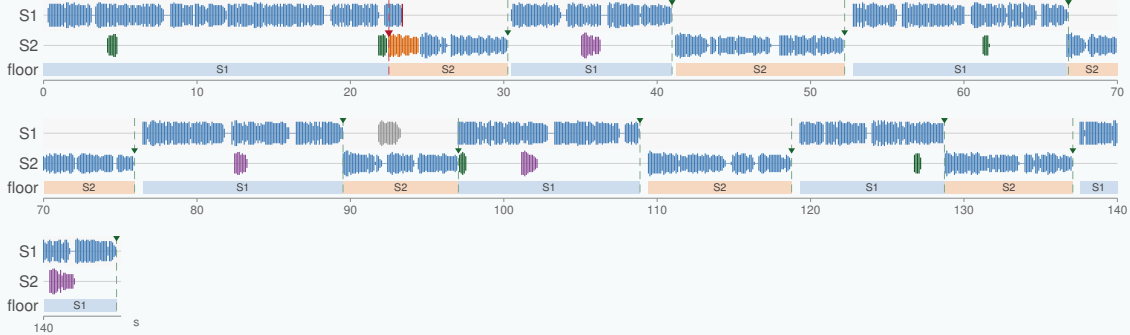
112.1–113.1 **S2** [laughter] Haha, yeah...

116.8–128.3 **S2** [turn] Uh, Mr. Patel in sophomore history, he'd give every war a soap opera plotline. He once acted out the Battle of Hastings with Playmobil figures. I think that's still the only reason I remember 1066. [▷ yields to S1](#)

128.6–141.5 **S1** [turn] That's so good. Honestly, I wish we had more teachers like that. Makes everything stick way better. Okay, so, like, for this exam, should we try giving each other questions but in funny voices or something? [▷ yields to S2](#)

141.6–147.8 **S2** [turn] Oh, uh, yeah, definitely. I bet I can come up with a really questionable British accent if you toss me a biology question. [▷ yields, end](#)

test split · tone: thoughtful and contemplative · noise: living room tv distant, 22.5 dB SNR; alarm beep; distant speech · voices: S1 F, S2 M
EOT 12 yields / 0 holds · INT 1 interruptions / 9 negatives (5 backchannels, 4 laughter) · 1 failed claims (excluded)

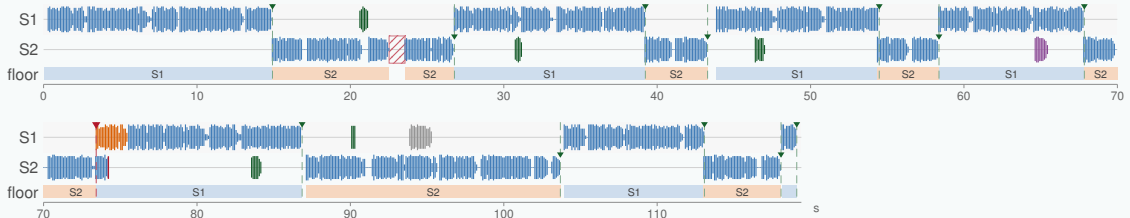


Transcript of test_00163

start—end (s) speaker [ground-truth label] text ▷ outcome

- 0.3–12.0 **S1** [turn] So, you know how I kept talking about redoing the kitchen? I finally went for it last year and honestly, it changed the whole vibe of the house. It was a headache at times but, oh my gosh, so worth it in the end.
- 4.2–4.9 **S2** [backchannel] No way!
- 12.0–23.4 **S1** [turn] Yeah! There was this one day, the plumbers accidentally drilled into the wrong wall and water started spraying everywhere. I actually stood there, just frozen, like, is this real life? Meanwhile, my dog's freaking out, slipping on the new tiles. ▷ cut off
- 21.8–22.4 **S2** [backchannel] Oh my—
- 22.5–24.5 **S2** [interruption] Wait, did the tiles survive all that? ▷ takes floor from S1
- 24.6–30.3 **S2** [turn] Because that's like, my biggest fear. Putting in these gorgeous floors and then—bam—disaster two days later. ▷ yields to S1
- 30.5–41.0 **S1** [turn] Totally! I was sure they'd be ruined, but they actually held up. Turns out ceramic is tougher than I thought. Plus, the dog learned not to run on wet floors, so, you know, bonus lesson. ▷ yields to S2
- 35.1–36.4 **S2** [laughter] Haha, smart dog.
- 41.2–52.2 **S2** [turn] I've been putting off starting on my own bathroom, um, mostly because I'm kind of terrified it'll just spiral out of control. Like, my friend tried to put up new tiles and ended up discovering mold behind the walls. ▷ yields to S1
- 52.8–66.8 **S1** [turn] Ugh, exactly! It always seems to snowball. With my kitchen, we found old wiring that literally looked like spaghetti. The contractor took one look and was just like, 'Nope, this is a fire hazard.' Delayed us two extra weeks. ▷ yields to S2
- 61.2–61.7 **S2** [backchannel] Mm-hm.
- 66.7–75.9 **S2** [turn] Yeah, that's my nightmare. My house is almost a hundred years old, so every time I poke around in a wall, I feel like I'm opening a portal to another era, and not in a cute way. ▷ yields to S1
- 76.5–89.5 **S1** [turn] Right? But honestly, I think it's those surprises that make you appreciate it more when it's finally done. I spent a week just standing in the kitchen, admiring my new cabinets. I didn't even cook much, just kept opening all the drawers. ▷ yields to S2
- 82.5–83.3 **S2** [laughter] Haha, yeah!
- 89.5–97.0 **S2** [turn] I do that too, whenever something's finally finished. I just walk around and touch things, like, 'Hey, this is mine now, I survived this ordeal!' ▷ yields to S1
- 91.8–93.3 **S1** [failed claim] Exactly! Oh, and get this—
- 97.0–108.9 **S1** [turn] —I tried to put in one of those fancy pull-out spice racks and somehow managed to install it upside down. I didn't even realize until all the little jars started rolling onto the floor. I just had to laugh. ▷ yields to S2
- 97.1–97.6 **S2** [backchannel] Yeah?
- 101.2–102.2 **S2** [laughter] Haha, oh man.
- 109.4–118.8 **S2** [turn] Reminds me of when I tried to replace a light fixture and ended up knocking out power to the whole upstairs. Had to call my dad in a panic. I swear, nothing ever goes as planned. ▷ yields to S1
- 119.3–128.7 **S1** [turn] Right, and yet we keep trying! I think that's half the fun though—having stories to tell. Plus, I feel like every mistake just, um, makes the place feel more like home, you know? ▷ yields to S2
- 126.8–127.2 **S2** [backchannel] Yeah.
- 128.8–137.1 **S2** [turn] Totally. Honestly, I kind of want to start on the bathroom now, just to see what stories I might end up with. Or at least, uh, some new tile patterns to obsess over. ▷ yields to S1
- 137.6–144.8 **S1** [turn] Do it! And when you inevitably mess something up, I'll bring snacks and moral support. Or, you know, help you mop up any surprise floods. ▷ yields, end
- 140.4–142.0 **S2** [laughter] Haha, I'll take you up on that.

test split · tone: thoughtful and contemplative · noise: quiet apartment, 31.1 dB SNR; alarm beep · voices: S1 F, S2 M
EOT 12 yields / 1 holds · INT 1 interruptions / 6 negatives (5 backchannels, 1 laughter) · 1 failed claims (excluded)



Transcript of test_00185

start—end (s) speaker [ground-truth label] text ▷ outcome

- 0.3–14.9 **S1** [turn] You know, I was thinking about our last park clean-up, and it's funny how those mornings always feel kind of impossible at first. Like, I woke up dreading the rain, but then by the time I got there, just seeing everyone with their gloves already on—it totally shifted my mood. ▷ yields to S2
- 14.9–22.5 **S2** [turn] Yeah, that's something about shared misery, right? There's this sense, uh, like we're all in it together, getting soaked but still laughing. I remember looking over and— ▷ holds, pauses
- 20.6–21.2 **S1** [backchannel] mm-hm
- 23.6–26.8 **S2** [turn] —seeing you trying to tape that sign up and it just kept falling down. ▷ yields to S1

26.8–39.2 **S1** [turn] Oh my gosh, yes! I was so determined that sign would stay put, and it just—well, it was like a comedy sketch. But I noticed all these strangers stepping in without a word, just holding the corners for me while I fumbled with the tape. **> yields to S2**

30.7–31.2 **S2** [backchannel] right

39.2–43.3 **S2** [turn] That's what gets me about these events—people just quietly pitching in, no fuss. **> yields to S1**

43.9–54.5 **S1** [turn] Exactly. And did you see the kid—maybe seven, eight years old—who started dragging that big trash bag half her size? Her dad looked so proud but also a little worried she'd tip over. **> yields to S2**

46.4–47.0 **S2** [backchannel] oh yeah

54.4–58.4 **S2** [turn] I was going to help, but she waved me off. Like, she wanted to do it herself, you know? **> yields to S1**

58.4–67.8 **S1** [turn] Yes—kids can be so determined. It reminded me of when I first volunteered, I was maybe her age, I thought picking up litter made me some sort of superhero for the day. **> yields to S2**

64.6–65.5 **S2** [laughter] haha right

67.8–74.2 **S2** [turn] That's probably why people keep coming back. Even for something as simple as moving mulch or stacking chairs, it feels, I guess, sort of heroic, in a— **> cut off**

73.4–75.5 **S1** [interruption] —in a small way, yeah, I get that. **> takes floor from S2**

75.6–86.9 **S1** [turn] I think there's something about everyone showing up, not for the recognition, just for the little things. Like, I met Marcy last month—she drove from two towns over just because she remembered the creek as a kid. **> yields to S2**

83.6–84.2 **S2** [backchannel] really?

87.1–95.7 **S2** [turn] That's beautiful. Makes me wonder how many of us have those little secret reasons. Sometimes when I'm raking leaves, I start thinking about my grandfather's garden and—

90.1–90.3 **S1** [backchannel] mm

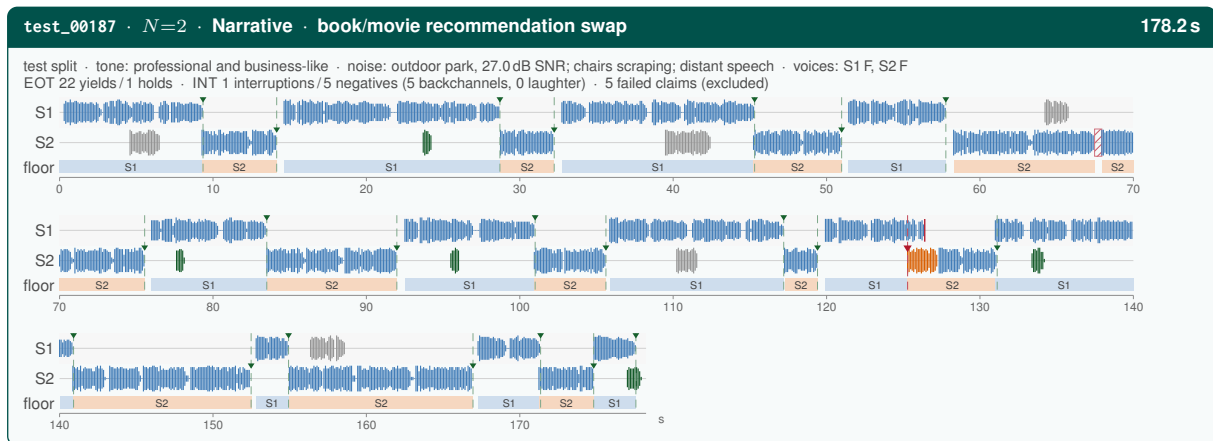
93.9–95.3 **S1** [failed claim] Sorry, can I ask—

95.8–103.7 **S2** [turn] —how he used to bring me out to plant tulips every spring. I think, uh, that's why I always sign up for planting over anything else. It just feels... familiar. **> yields to S1**

103.9–113.1 **S1** [turn] That makes so much sense. I love how these events turn into stories—like, everyone comes with a memory, and somehow we all leave with a new one. Even when it rains. **> yields to S2**

113.0–118.1 **S2** [turn] Yeah, yeah. Honestly, I look forward to whatever the weather throws at us. It's part of the story now. **> yields to S1**

118.1–119.1 **S1** [turn] Absolutely. **> yields, end**



Transcript of test_00187

start—end (s) speaker [ground-truth label] text **> outcome**

0.3–9.3 **S1** [turn] So, you said you were looking for something to read on your trip. I just finished "Klara and the Sun," and I think it might be up your alley. It's efficient, almost clinical, but deeply moving underneath. **> yields to S2**

4.6–6.6 **S2** [failed claim] Oh, Kazuo Ishiguro, right?

9.3–14.2 **S2** [turn] I've noticed his books tend to sneak up on you emotionally. What about this one stood out? **> yields to S1**

14.6–28.7 **S1** [turn] Yeah, exactly. Well, Klara's perspective is so stripped down, but it exposes all these—um, vulnerabilities in the human characters. The way she observes loneliness in such an analytical way, it made me think about people at work, honestly. You know, how much we mask. **> yields to S2**

23.7–24.3 **S2** [backchannel] Right.

28.7–32.2 **S2** [turn] That sounds a bit like "Never Let Me Go" in terms of the subtlety. **> yields to S1**

32.7–45.3 **S1** [turn] It is, but less bleak somehow. There's this persistent hopefulness, or maybe just acceptance, which is refreshing. I actually finished it in one weekend. I found myself thinking about the office dynamic—especially after that reorg last quarter. **> yields to S2**

39.5–42.4 **S2** [failed claim] Yeah, I think we've all been a bit Klara-ish since the reorg.

45.2–51.0 **S2** [turn] So for movies, I actually watched "The Farewell" last night. Not recent, but it was new to me. Have you seen it? **> yields to S1**

51.4–54.6 **S1** [turn] Oh, Awkwafina, right? Haven't gotten around to it yet.

54.6–57.8 **S1** [turn] Isn't it about a family hiding an illness from their grandmother? **> yields to S2**

58.3–67.5 **S2** [turn] Yeah, exactly. What really struck me was how, um, everyone coped by just pretending. My family does something similar, but with work stress instead of illness— **> holds, pauses**

64.2–65.8 **S1** [failed claim] We all perform, don't we—

67.9–75.5 **S2** [turn] —right, and no one ever says directly what the problem is. It's handled in small, indirect ways. I found that comforting, actually. **> yields to S1**

76.0–83.5 **S1** [turn] Interesting. I usually get frustrated by characters who don't just say what they mean, but maybe that's more realistic than I'd like to admit. **> yields to S2**

77.6–78.1 **S2** [backchannel] Mm-hm.

83.5–92.0 **S2** [turn] I used to think directness was always best, but after this past year, I'm not sure. There's something to be said for kindness, even when it's a bit of a lie. **> yields to S1**

92.5–101.0 **S1** [turn] That makes sense. Actually, that's part of what resonated in "Klara and the Sun", too. It's all these little negotiations between truth and what people need to hear. **> yields to S2**

95.5–96.1 **S2** [backchannel] Yeah.

100.9–105.6 **S2** [turn] So did you ever have a book or film actually change how you handled something at work? **> yields to S1**

105.9–117.2 **S1** [turn] Not dramatically, but after reading "Station Eleven," I did start keeping a little notebook on my desk. Just, you know, to jot down things that matter before they slip away. It helped during all those project transitions. **> yields to S2**

110.2–111.6 **S2** [failed claim] That's a good idea.

117.3–119.4 **S2** [turn] I might steal that scheme, honestly. **> yields to S1**

119.9–126.4 **S1** [turn] By all means. Actually, if you want something light next, "The Thursday Murder Club" was surprisingly smart—like, quick, but not mindless. **> cut off**

125.3–127.3 **S2** [interruption] I've seen it everywhere! Worth the hype? ▷ takes floor from S1

127.3–131.1 **S2** [turn] Because I've been putting it off, thinking it's just cozy mystery fluff. ▷ yields to S1

131.0–140.9 **S1** [turn] I mean, it's a cozy, but it's got teeth. The characters are sharp; the humor's understated. Honestly, good for a flight, especially if you want something that doesn't drain you but isn't, well, shallow. ▷ yields to S2

133.4–134.2 **S2** [backchannel] Perfect.

140.9–152.5 **S2** [turn] Okay, top choice for you: "Columbus." It's a movie, but not your typical one. Minimal dialogue, but every scene feels intentional. I watched it on a layover and ended up scribbling half a notebook of thoughts. ▷ yields to S1

152.8–154.9 **S1** [turn] That's the one with John Cho, right? ▷ yields to S2

155.0–157.2 **S2** [turn] Yeah, exactly. It's so quiet but—

156.4–158.7 **S1** [failed claim] Sometimes I need quiet, actually.

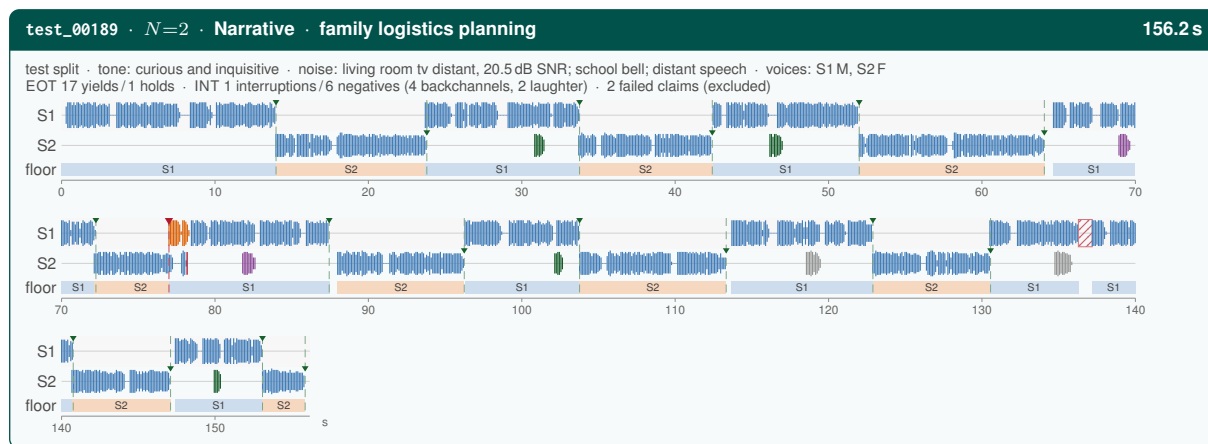
157.2–167.0 **S2** [turn] —the architecture becomes a character. It gave me a new way to look at spaces, even around the office. Like, I started noticing the way light moves in at different times of day. ▷ yields to S1

167.3–171.4 **S1** [turn] Sounds like something I'd appreciate. I'll queue it up for my own next night in. ▷ yields to S2

171.2–174.8 **S2** [turn] Good. And I'll try "Klara and the Sun" for my flight. Thanks for the tip. ▷ yields to S1

174.8–177.6 **S1** [turn] Let me know if it lands for you as much as it did for me. ▷ yields, end

177.0–177.9 **S2** [backchannel] Will do.



Transcript of test_00189

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–14.0 **S1** [turn] So, I was thinking about how we handled last year's road trip, and honestly, I'm still amazed we managed to get everyone out the door before eight. That almost never happens, especially with your dad insisting that we all eat a real breakfast first. ▷ yields to S2

14.0–23.8 **S2** [turn] Oh my gosh, yeah! That morning was absolutely chaos. I thought we were gonna lose Aunt Linda when she couldn't find her sunglasses and started emptying her purse in the driveway. ▷ yields to S1

23.7–33.8 **S1** [turn] And then Sam was still in pajamas—he just walked out, sleepy-faced, like nothing was wrong. Meanwhile, I was counting suitcases, thinking, okay, did we forget the cooler this time? ▷ yields to S2

30.8–31.5 **S2** [backchannel] Right, right!

33.7–42.4 **S2** [turn] It's always the cooler! For some reason, someone always puts it right by the front door, but then it just sits there while everyone thinks someone else already loaded it. ▷ yields to S1

42.5–52.0 **S1** [turn] Exactly. I've started making a checklist on my phone because of that. Otherwise, we end up an hour into the drive and realize we have twelve sandwiches but no drinks. ▷ yields to S2

46.2–47.0 **S2** [backchannel] Yeah, totally.

52.0–64.1 **S2** [turn] Um, do you remember the time when we tried to organize a group breakfast before the graduation ceremony? Somehow, half the family ended up at that weird pancake place and the rest of us were at the diner across town. ▷ yields to S1

64.7–72.2 **S1** [turn] Oh, I remember. The texts were flying—"Where are you?" "We're at the blue chairs!" Meanwhile, we were sitting under a stuffed moose head, totally alone. ▷ yields to S2

68.9–69.7 **S2** [laughter] Haha, yes!

72.1–78.2 **S2** [turn] I think that's why I get so intense about sending everyone a Google Calendar link these days. But, like, even then, someone always—▷ cut off

77.0–78.3 **S1** [interruption] Doesn't look at it. Yeah. ▷ takes floor from S2

78.5–87.4 **S1** [turn] Last Thanksgiving, I set up a shared doc with everyone's arrival times, and Uncle Mark still called me at ten the night before, asking if he should bring dessert. ▷ yields to S2

81.8–82.7 **S2** [laughter] Haha, oh Mark.

88.0–96.2 **S2** [turn] Honestly, I think he just likes having an excuse to call. But, wait—were you the one who coordinated the secret Santa that year or was that Mom? ▷ yields to S1

96.3–103.8 **S1** [turn] That was actually me, but Mom did all the reminding. I just set everyone up with names; she chased people down. I think she enjoyed that part. ▷ yields to S2

102.2–102.7 **S2** [backchannel] Uh-huh.

103.8–113.3 **S2** [turn] I wish I had her patience. If I have to explain one more time how to use the group chat, I'm going to lose it. Remember how Grandpa sends every reply as a new text? ▷ yields to S1

113.7–122.9 **S1** [turn] Yeah, and none of them actually land in the right thread. One year, I ended up with three separate 'bring rolls!' messages, all from him, all in different conversations. ▷ yields to S2

118.6–119.5 **S2** [failed claim] Oh my gosh, yes!

122.9–130.6 **S2** [turn] We definitely need a new system for this summer's reunion. Maybe something foolproof—like printed itineraries taped to everyone's door. ▷ yields to S1

130.5–136.3 **S1** [turn] Haha, yeah, old-school! Or maybe we need a bullhorn and just announce departure times every thirty minutes. ▷ holds, pauses

134.8–135.9 **S2** [failed claim] Wait, but seriously—

137.2–140.8 **S1** [turn] —I mean, could work! But, yeah, if you've got a better plan, I'm all ears. ▷ yields to S2

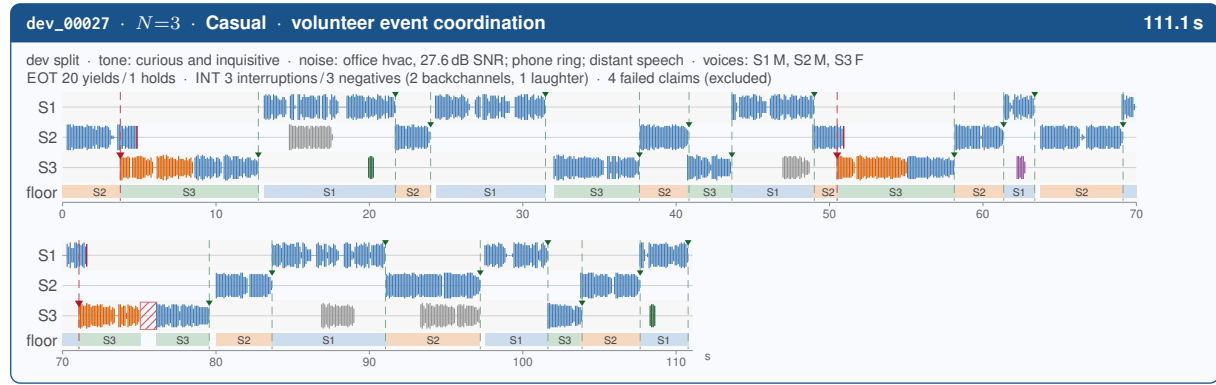
140.7–147.1 **S2** [turn] Well, what if we assign everyone one person to keep tabs on? Like buddy system, but for adults who might wander off. ▷ yields to S1

147.4–153.1 **S1** [turn] That actually might be brilliant. Everyone gets a 'buddy'—no one left behind at the rest stop ever again. ▷ yields to S2

150.0–150.4 **S2** [backchannel] Mm-hm!

153.1–155.9 **S2** [turn] Alright, I'm putting that on my planning spreadsheet right now. ▷ yields, end

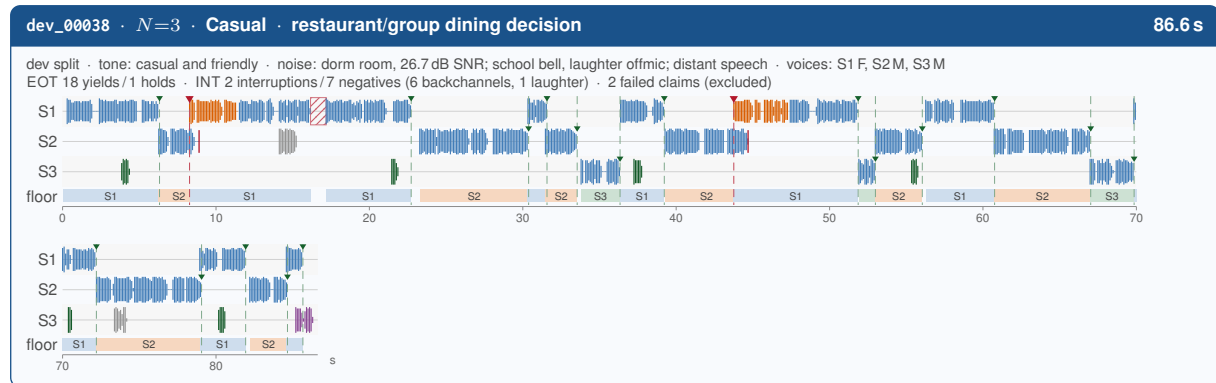
K.2 Conversations with $N = 3$ speakers



Transcript of dev_00027

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–4.9 **S2** [turn] Alright, so—uh, are we all set for the park cleanup this weekend, or are there still things we need to nail down? ▷ cut off
3.8–8.5 **S3** [interruption] Actually, wait, do we even have enough people? I feel like last time it was just, like, six of us. ▷ takes floor from S2
8.6–12.8 **S3** [turn] Because, honestly, that was brutal. I don't want to haul bags for three hours again. ▷ yields to S1
13.2–16.8 **S1** [turn] We have eight signups so, you know, that's technically more than six.
14.8–17.6 **S2** [failed claim] Right, but are three of those just Sarah's friends who never show up?
16.9–21.7 **S1** [turn] —Yeah, possibly. I wouldn't bet money on them appearing, but hey, optimism. ▷ yields to S2
20.0–20.4 **S3** [backchannel] mm-hm
21.7–24.0 **S2** [turn] Haha, yeah, Sarah's friends exist in theory. ▷ yields to S1
24.3–31.5 **S1** [turn] I've literally never spoken to any of them. So, um, what about supplies? Do we even have enough gloves this time? ▷ yields to S3
32.0–37.6 **S3** [turn] I grabbed a box from the rec center but, I mean, they're that weird size? Like, only fit if you have tiny hands. ▷ yields to S2
37.7–40.8 **S2** [turn] Oh, I can pick up some extra ones. Do we need trash bags too, or...? ▷ yields to S3
40.8–43.6 **S3** [turn] Definitely. I went through, like, four last time on my own. ▷ yields to S1
43.6–49.0 **S1** [turn] Yeah. And maybe those grabber things, so I don't have to touch whatever that sticky mess was. ▷ yields to S2
47.0–48.7 **S3** [failed claim] Ew, yeah, that was horrifying.
48.9–50.9 **S2** [turn] I think I saw them on sale at the hardware store. Should I grab a couple? ▷ cut off
50.5–55.0 **S3** [interruption] How many is a couple though? Like, actually two, or 'a couple' as in, like, everyone gets one? ▷ takes floor from S2
55.1–58.1 **S3** [turn] Because if I'm using my hands I'm bailing after lunch. ▷ yields to S2
58.2–61.4 **S2** [turn] No, like, I'll get six. Or maybe eight. Worst case, we fight over them. ▷ yields to S1
61.4–63.4 **S1** [turn] Winner gets to pick music too. ▷ yields to S2
62.2–62.8 **S3** [laughter] haha
63.8–69.1 **S2** [turn] Oh—speaking of music, did anyone bring a speaker last time? I feel like it was just my phone and that was... not... loud. ▷ yields to S1
69.1–71.6 **S1** [turn] That was you. It kept cutting out every time a car went by. ▷ cut off
71.1–75.1 **S3** [interruption] Yeah, can someone please bring an actual speaker? Not just a phone in a cup? ▷ takes floor from S1 ▷ holds, pauses
76.1–79.6 **S3** [turn] I have one, just remind me Friday so I don't forget it at home again. ▷ yields to S2
80.0–83.7 **S2** [turn] Will do. I'll text you. Uh, what about snacks? Are we thinking bagels again or...? ▷ yields to S1
83.7–91.0 **S1** [turn] I'm still recovering from those last bagels. Maybe fruit this time. Or, like, something that doesn't get gluey in your mouth. ▷ yields to S2
86.9–89.0 **S3** [failed claim] orange slices are superior anyway
91.1–94.8 **S2** [turn] Mmm, oranges are good. I'll bring a couple bags. Maybe some granola bars too?
93.3–97.2 **S3** [failed claim] Wait, are those the kind you brought last time, the ones that crumble instantly?
94.9–97.2 **S2** [turn] —No, I'll get the chewy ones, not the oat brick ones. ▷ yields to S1
97.5–101.6 **S1** [turn] Chewy is acceptable. Also less likely to attract pigeons. ▷ yields to S3
101.6–103.9 **S3** [turn] That's a lie. Everything attracts pigeons. ▷ yields to S2
103.8–107.7 **S2** [turn] Haha, I'll bring napkins just in case. Sounds like we sort of have a plan now? ▷ yields to S1
107.6–110.8 **S1** [turn] Yeah, that's probably as organized as we ever get. ▷ yields, end
108.3–108.7 **S3** [backchannel] mm-hm



Transcript of dev_00038

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.3 **S1** [turn] Alright, so we have to pick something—I'm not letting us stand out here for another half hour, my feet are killing me. ▷ yields to S2
3.9–4.4 **S3** [backchannel] Same.
6.3–8.9 **S2** [turn] Agreed. But let's be strategic, like, do you guys want something heavy or— ▷ cut off
8.3–11.4 **S1** [interruption] Light! No way am I doing pasta right now, not with this heat. ▷ takes floor from S2

11.5–16.2 **S1** [turn] Seriously, something with a salad, or like, those poke bowls from that place up the block? ▷ holds, pauses

14.1–15.3 **S2** [failed claim] Oh, actually—

17.2–22.7 **S1** [turn] —unless you two are absolutely dying for pizza, in which case, I'll survive. But poke is my vote. ▷ yields to S2

21.5–21.9 **S3** [backchannel] Mm-hm.

23.3–30.4 **S2** [turn] I was just gonna say that poke spot is actually closed today, I checked earlier. Their sign said staff meeting or something. ▷ yields to S1

30.3–31.6 **S1** [turn] Ugh, of course it is. ▷ yields to S2

31.5–33.5 **S2** [turn] Yeah. Sushi next door's open though. ▷ yields to S3

33.8–36.3 **S3** [turn] That place is okay. Kinda slow. ▷ yields to S1

36.4–39.2 **S1** [turn] Yeah but I mean, it's air conditioned. I'm melting. ▷ yields to S2

37.2–37.8 **S3** [backchannel] Valid.

39.3–44.7 **S2** [turn] It's true. The other option is Indian, but that'll be spicier and probably not light. ▷ cut off

43.8–47.4 **S1** [interruption] Too much for me right now—wait, what about that falafel place? ▷ takes floor from S2

47.4–51.9 **S1** [turn] You know, the one with the patio? I know it's not super exciting but at least it's light. ▷ yields to S3

51.9–53.0 **S3** [turn] Has shade though. ▷ yields to S2

53.0–56.0 **S2** [turn] And we could probably get a table, it's never packed. ▷ yields to S1

55.3–55.8 **S3** [backchannel] Yeah.

56.3–60.8 **S1** [turn] Okay, are we actually landing there or are we gonna circle for another twenty minutes? ▷ yields to S2

60.7–67.0 **S2** [turn] I'm good with it. Falafel's easy. Or, unless, Rob, do you actually want pizza? You haven't said much. ▷ yields to S3

67.0–69.9 **S3** [turn] Nah, pizza's fine but it'll put me to sleep. ▷ yields to S1

69.8–72.2 **S1** [turn] Exactly! I knew someone was going to say that. ▷ yields to S2

70.4–70.7 **S3** [backchannel] Mm.

72.2–74.6 **S2** [turn] Ha, alright. Falafel it is then?

73.4–74.2 **S3** [failed claim] Wait, but—

74.7–79.1 **S2** [turn] —unless there's a last second protest. Speak now or get in line. ▷ yields to S1

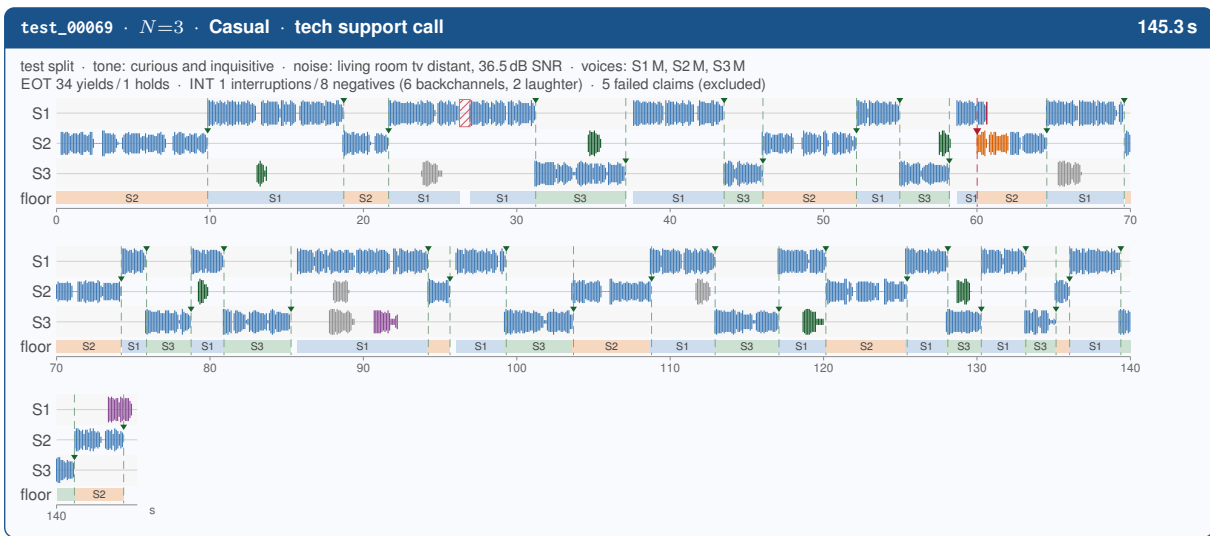
78.9–81.9 **S1** [turn] Nope, falafel, done—let's go before I overthink it. ▷ yields to S2

80.2–80.7 **S3** [backchannel] Sure.

82.2–84.7 **S2** [turn] Decision achieved. Miraculous. ▷ yields to S1

84.6–85.7 **S1** [turn] Don't get used to it. ▷ yields, end

85.2–86.3 **S3** [laughter] Haha, noted.



Transcript of test_00069

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.8 **S2** [turn] Hey guys, thanks for hopping on the call. Uh, so Mark, you mentioned your laptop keeps freezing when you log in? Can you walk us through what exactly's happening? ▷ yields to S1

9.9–18.7 **S1** [turn] Yeah, um, so it kind of gets stuck right after I put in my password. Like, the spinning wheel just, uh, keeps spinning? And nothing loads for like, I dunno, five minutes or something. ▷ yields to S2

13.1–13.7 **S3** [backchannel] Mm-hmm.

18.7–21.6 **S2** [turn] Right, so it doesn't get to the desktop, just hangs there. ▷ yields to S1

21.7–26.3 **S1** [turn] Yeah. The weird thing is, I can hear like, uh, fans turning on? And then it gets really hot. ▷ holds, pauses

23.8–25.2 **S3** [failed claim] Oh that's not good—

27.0–31.2 **S1** [turn] —so I usually, uh, just force shut it down, but then it's the same when I restart. ▷ yields to S3

31.2–37.1 **S3** [turn] Yeah, probably a process stuck at boot or something. Have you noticed any weird pop-ups, like, uh, before this started? ▷ yields to S1

34.7–35.6 **S2** [backchannel] Good question.

37.6–43.5 **S1** [turn] Not really? I mean, maybe one or two like—'your antivirus needs attention' things? But I usually just close those. ▷ yields to S3

43.5–46.0 **S3** [turn] Yeah, classic. Sometimes that's all it takes. ▷ yields to S2

46.0–52.1 **S2** [turn] Have you tried booting in safe mode yet? That can sometimes, uh, get around startup issues. ▷ yields to S1

52.2–55.0 **S1** [turn] No, is that—uh, is that like hitting F8, or...? ▷ yields to S3

55.0–58.2 **S3** [turn] Yeah, F8 or hold shift when you restart on newer ones. ▷ yields to S1

57.5–58.3 **S2** [backchannel] Exactly.

58.7–60.6 **S1** [turn] Okay, I can try that. Should I like— ▷ cut off

60.0–62.1 **S2** [interruption] If you want, we can walk you through it now? ▷ takes floor from S1

62.2–64.6 **S2** [turn] Just, uh, do you have the laptop in front of you? ▷ yields to S1

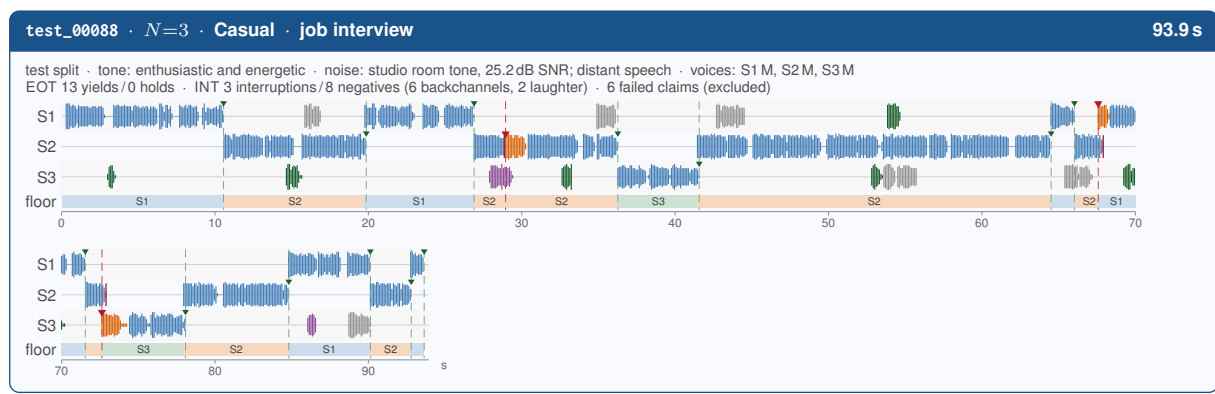
64.5–66.6 **S1** [turn] Yeah, I've got it right here, let me just—

65.3–66.8 **S3** [failed claim] No worries, take your time.

66.7–69.6 **S1** [turn] Okay, so do I shut it down again? Or is it—uh— ▷ yields to S2

69.7–74.2 **S2** [turn] Yeah, let's do a restart and this time just hold shift as soon as it's booting up. ▷ yields to S1

74.3–75.9 **S1** [turn] Alright, restarting now. ▷ yields to S3
 75.8–78.8 **S3** [turn] It should give you a blue menu, says something like 'troubleshoot.' ▷ yields to S1
 78.8–80.9 **S1** [turn] Oh, yeah! I see it, uh, 'choose an option.' ▷ yields to S3
 79.2–79.9 **S2** [backchannel] Perfect.
 80.9–85.3 **S3** [turn] Now just hit 'troubleshoot,' then 'advanced options,' and safe mode should be in there. ▷ yields to S1
 85.7–89.4 **S1** [turn] Okay, okay, um... I'm clicking through. Sorry if I'm kinda slow at this.
 87.8–89.4 **S3** [failed claim] Yeah, take your time. You're good.
 88.1–89.1 **S2** [failed claim] No worries at all.
 89.6–94.2 **S1** [turn] Alright, I think it's going into safe mode now. Uh... oh wow, it's actually showing the desktop. ▷ yields to S2
 90.7–92.3 **S3** [laughter] Aha, there we go!
 94.2–95.6 **S2** [turn] That already tells us a lot. ▷ yields to S1
 96.0–99.3 **S1** [turn] So is it, uh, like a virus or just a startup thing? Or— ▷ yields to S3
 99.2–103.7 **S3** [turn] Probably some program misbehaving at startup. Could be the antivirus, honestly. ▷ yields to S2
 103.6–108.8 **S2** [turn] Yeah, let's check what runs at startup. Open task manager and go to the startup tab. ▷ yields to S1
 108.7–112.9 **S1** [turn] Okay, um... wow, there's like a bunch here. Should I just, uh, disable everything? ▷ yields to S3
 111.7–112.7 **S2** [failed claim] Not everything—
 112.9–117.1 **S3** [turn] —just the weird ones, like stuff you don't recognize. Leave Windows things alone. ▷ yields to S1
 117.0–120.2 **S1** [turn] Alright, like, 'HP Support Assistant'? I can turn that off? ▷ yields to S2
 118.7–120.0 **S3** [backchannel] Yeah, that's fine.
 120.2–125.4 **S2** [turn] Yep, and anything with 'Update' in it that's not Windows, probably safe to disable. ▷ yields to S1
 125.3–128.1 **S1** [turn] Alright, done. Should I just restart like, normal now? ▷ yields to S3
 128.0–130.3 **S3** [turn] Yeah, see if it gets past that spinning wheel. ▷ yields to S1
 128.7–129.6 **S2** [backchannel] Fingers crossed.
 130.3–133.2 **S1** [turn] Alright, rebooting now. Hey, it looks like it's working! ▷ yields to S3
 133.1–135.2 **S3** [turn] Boom—look at that. Fixed. ▷ yields to S2
 135.1–136.0 **S2** [turn] Nice work, Mark. ▷ yields to S1
 136.1–139.4 **S1** [turn] Thanks, guys. Uh, sorry for all the questions, but I appreciate it. ▷ yields to S3
 139.3–141.2 **S3** [turn] No worries. That's what we're here for. ▷ yields to S2
 141.2–144.4 **S2** [turn] Yeah, and if this pops up again, you know where to find us. ▷ yields, end
 143.4–145.0 **S1** [laughter] Haha, yeah, thanks again!

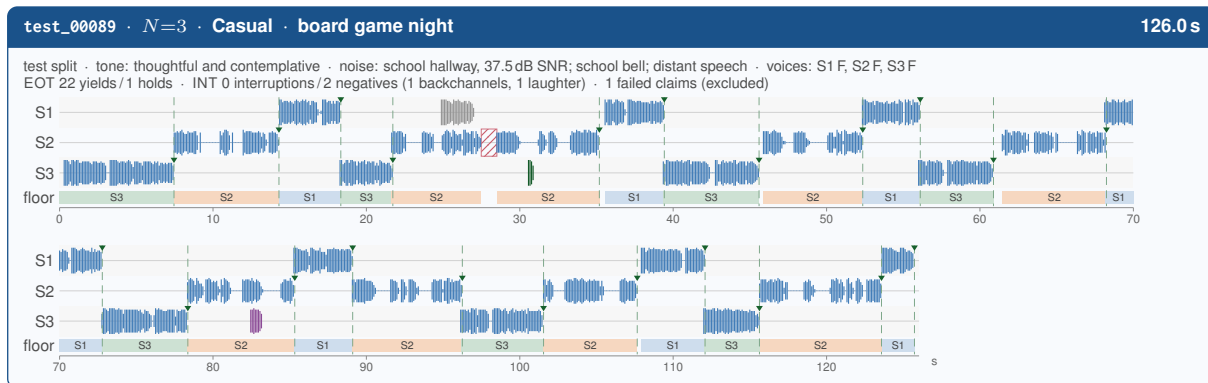


Transcript of test_00088

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–10.6 **S1** [turn] Alright, Ben, thanks so much for joining us today! We're really just gonna chat, nothing too serious, so, uh, first—could you just tell us, like, what drew you to this role? ▷ yields to S2
 3.0–3.5 **S3** [backchannel] yeah
 10.6–19.9 **S2** [turn] Yeah, um, honestly what caught my eye was just how much, uh, freedom it seemed like the team has. I love when folks can, you know, experiment a bit instead of, like, always following a checklist. ▷ yields to S1
 14.7–15.7 **S3** [backchannel] Right, right.
 15.9–16.9 **S1** [failed claim] Mhm, yeah, for sure.
 19.8–26.9 **S1** [turn] So you uh—you like the sort of, messy creative side? Or is it more, like, you just don't like being told what to do? ▷ yields to S2
 26.9–28.9 **S2** [turn] Haha, well, maybe a little bit of both? ▷ cut off
 27.9–29.4 **S3** [laughter] Haha yeah, I hear that.
 28.9–30.3 **S2** [interruption] But no, seriously— ▷ takes floor from S1
 30.4–36.3 **S2** [turn] —I think you get better stuff out of people when you let them try things out. Not just me—like, teams in general. ▷ yields to S3
 32.6–33.3 **S3** [backchannel] Agreed.
 34.9–36.2 **S1** [failed claim] That's totally our vibe.
 36.3–41.6 **S3** [turn] Do you, uh, do you have, like, an example? Something where you, I dunno, tried something out and it worked or... didn't? ▷ yields to S2
 41.5–43.4 **S2** [turn] Oh, oh yeah. Um, at my last job I, uh—
 42.7–44.6 **S1** [failed claim] Wait, was that the museum thing or—
 43.5–56.5 **S2** [turn] —yeah, at the science center. We had this project to uh, redesign the ticket flow, and the old process took, like, ages. So, uh, I sketched up a super basic prototype, like, on paper? And the leads actually just let me show it to people at the front desk.
 52.8–53.5 **S3** [backchannel] Nice.
 53.6–55.8 **S3** [failed claim] And what happened? Was it a disaster or—
 53.8–54.7 **S1** [backchannel] That's awesome.
 56.6–64.5 **S2** [turn] —people actually loved it! I mean, it wasn't perfect, but uh, we cut wait times by, like, twelve minutes? Management didn't believe the numbers at first, haha. ▷ yields to S1
 64.5–66.0 **S1** [turn] Wow, that's a huge difference. ▷ yields to S2
 65.4–67.3 **S3** [failed claim] Yeah, yeah that's, uh, impressive.
 66.1–67.9 **S2** [turn] Well, uh, thanks! Got lucky, I think. ▷ cut off
 67.6–68.3 **S1** [interruption] I dunno— ▷ takes floor from S2
 68.4–71.5 **S1** [turn] —you make your own luck sometimes, you know? That's initiative. ▷ yields to S2
 69.2–70.3 **S3** [backchannel] Exactly.

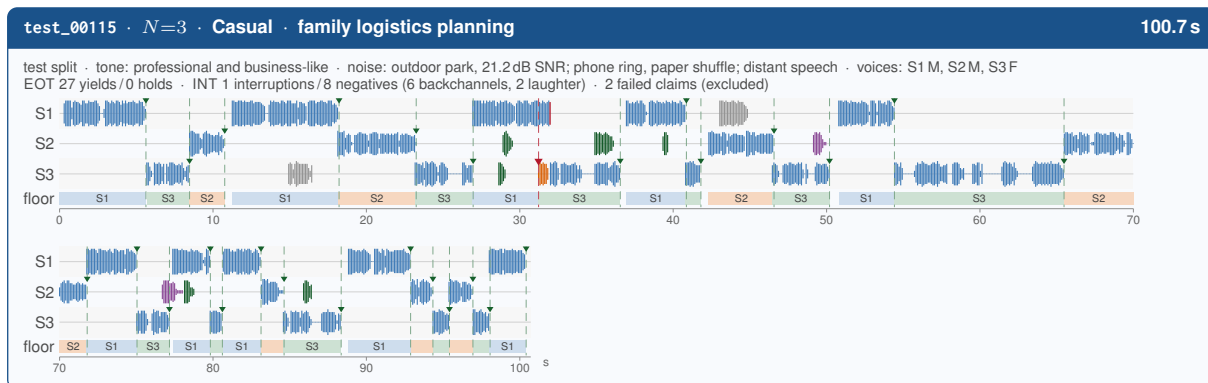
71.6–72.9 **S2** [turn] I appreciate that. > cut off
 72.6–74.3 **S3** [interruption] So—let me just ask— > takes floor from S2
 74.4–78.1 **S3** [turn] —do you get bored easily? Or like, what keeps things exciting for you at work? > yields to S2
 78.0–84.8 **S2** [turn] Oh wow, um, yeah—I do kinda need a challenge. If I can mix things up, or like, teach somebody else what I’ve learned, that sort of lights me up, you know? > yields to S1
 84.8–86.7 **S1** [turn] That totally comes through, honestly.
 86.1–86.6 **S3** [laughter] haha
 86.8–90.1 **S1** [turn] Okay, cool, this is great—I feel like I know you already. > yields to S2
 88.7–90.2 **S3** [failed claim] Yeah, honestly, same here.
 90.1–92.8 **S2** [turn] This has actually been kinda fun, not gonna lie. > yields to S1
 92.8–93.6 **S1** [turn] Awesome to hear! > yields, end



Transcript of test_00089

start–end (s) speaker [ground-truth label] text > outcome

0.3–7.5 **S3** [turn] Alright, so—are we really doing Codenames again? Not that I’m complaining, I just feel like we play it every single time we hang out. > yields to S2
 7.5–14.3 **S2** [turn] It’s tradition at this point. If we don’t, does the universe collapse or something? > yields to S1
 14.3–18.3 **S1** [turn] I mean, there’s something comforting about having a favorite. It’s like warm soup, you know? > yields to S3
 18.3–21.7 **S3** [turn] Yes! Except, uh, it’s competitive soup, and I keep burning my tongue. > yields to S2
 21.6–27.5 **S2** [turn] You can switch it up. There’s, like, a pile of games literally right there— > holds, pauses
 24.9–27.0 **S1** [failed claim] But none of those ever actually get played—
 28.5–35.2 **S2** [turn] —and we still end up here. Yeah, I know. It’s a cardboard decor collection. > yields to S1
 30.6–30.9 **S3** [backchannel] mm-hm
 35.6–39.4 **S1** [turn] It’s honestly nice, though. I like that we’re predictable in this one tiny way. > yields to S3
 39.4–45.6 **S3** [turn] Okay but can we admit, the red team always wins. Like, statistically impossible, but it still happens. > yields to S2
 45.9–52.4 **S2** [turn] I’d argue correlation, not causation. Maybe you’re just always on red. > yields to S1
 52.3–56.1 **S1** [turn] Or maybe we’re subconsciously rooting for you to win, Claire. It’s very possible. > yields to S3
 56.0–60.9 **S3** [turn] Aww, conspiracy confirmed! Friends secretly sabotage themselves for my confidence. > yields to S2
 61.5–68.2 **S2** [turn] If only board game night could fix my confidence. Do you think Avalon would help? > yields to S1
 68.1–72.8 **S1** [turn] Maybe if I could ever remember how that one works. I always forget the rules halfway through. > yields to S3
 72.8–78.4 **S3** [turn] Same! Like every game with traitors, my brain just panics. I can’t even keep my poker face together. > yields to S2
 78.3–85.3 **S2** [turn] Your poker face is basically you giggling and looking at the ceiling. Super subtle. > yields to S1
 82.5–83.2 **S3** [laughter] Haha, rude.
 85.3–89.1 **S1** [turn] I think it’s endearing. It means we trust you. Or just that you’re a terrible liar. > yields to S2
 89.1–96.3 **S2** [turn] Oh she’s the worst liar. But it’s, uh, charming? In a completely unhelpful way. > yields to S3
 96.1–101.5 **S3** [turn] I mean, it’s not like you’re any better. Remember when you tried to convince us you’re not the spy and you just— > yields to S2
 101.5–107.7 **S2** [turn] Yeah, yeah, I said my own code word by accident. The shame lives on. > yields to S1
 107.9–112.1 **S1** [turn] Honestly, I kind of like that we’re all so bad at bluffing. It’s more fun that way. > yields to S3
 112.0–115.6 **S3** [turn] Right? If we ever played with real bluffers, we’d lose in, like, two minutes. > yields to S2
 115.6–123.6 **S2** [turn] So we’ll just stick to Codenames. It’s safe. No bluffing, just, uh, miscommunication. > yields to S1
 123.6–125.7 **S1** [turn] That’s our specialty. Alright, teams? > yields, end



Transcript of test_00115

start–end (s) speaker [ground-truth label] text > outcome

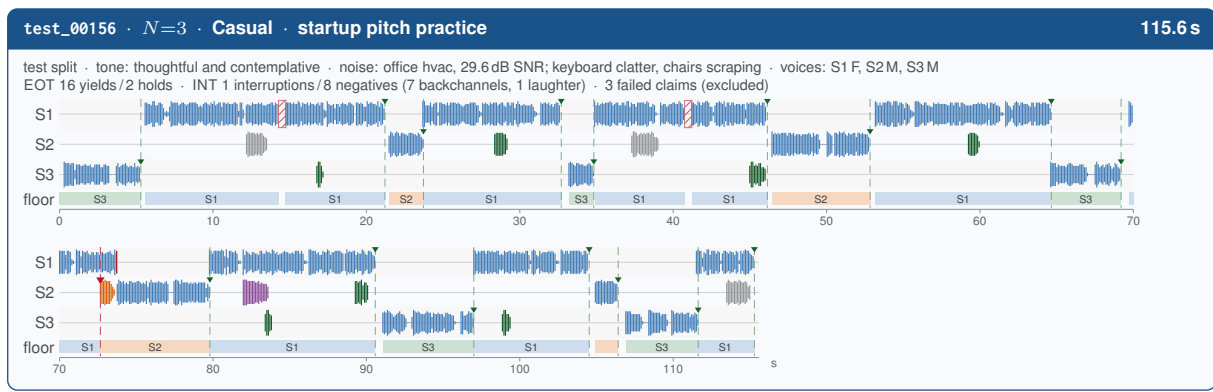
0.3–5.6 **S1** [turn] So uh... I guess we should start with Saturday, right? Who’s actually—um, who’s actually around in the morning? > yields to S3

test_00143 · $N=3$ · Casual · roommate household planning 84.5 s



51.9–56.7 **S2[turn]** The Roomba like, got stuck under the couch yesterday. I rescued it but it was looking kinda. . . sad. **▷ yields to S1**

53.2–54.1 **S3** [backchannel] Poor thing.
 56.7–57.9 **S1** [turn] Nooo, my child! ▷ yields to S3
 58.0–60.4 **S3** [turn] Haha, we need to tape googly eyes on it again.
 58.8–60.0 **S2** [backchannel] Yeah definitely.
 60.6–64.5 **S3** [turn] Um, switching topics for a sec, has anyone figured out the grocery run yet? ▷ yields to S1
 64.4–70.7 **S1** [turn] I was gonna ask! Because we're, like, out of cereal and coffee and, uh, basically everything that keeps me alive in the morning. ▷ yields to S2
 70.6–74.5 **S2** [turn] I can go tomorrow if someone, uh, gives me a list. I'll forget half without it. ▷ yields to S3
 74.4–79.0 **S3** [turn] I'll text you one! But only if you promise not to buy those weird off-brand Oreos again. ▷ yields to S2
 79.0–80.2 **S2** [turn] No promises. ▷ yields to S1
 80.7–82.2 **S1** [turn] Oh come on, Sam! ▷ yields to S3
 82.1–84.2 **S3** [turn] Haha, alright, call this meeting adjourned? ▷ yields, end
 82.6–83.4 **S2** [backchannel] Agreed.

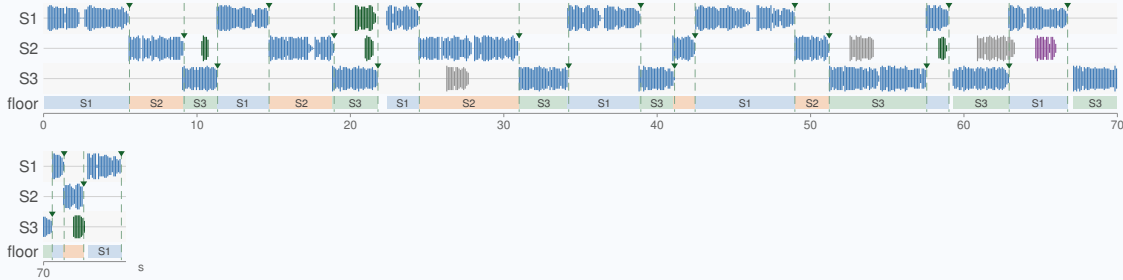


Transcript of test_00156

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–5.3 **S3** [turn] Alright, uh, whenever you're ready, Anna, just start your pitch. Take your time—no rush. ▷ yields to S1
 5.6–14.3 **S1** [turn] Okay, thanks, Mark. So, um, imagine an app that helps people manage their mental health by connecting them with, like, tiny daily check-ins, sort of like
 a—▷ holds, pauses
 12.2–13.5 **S2** [failed claim] Like a virtual buddy?
 14.8–21.2 **S1** [turn] —yeah, almost like a virtual buddy, exactly. But it's more personalized, right? It adapts based on mood tracking. ▷ yields to S2
 16.8–17.2 **S3** [backchannel] mm-hm
 21.5–23.7 **S2** [turn] So, it's not a therapist replacement, right? ▷ yields to S1
 23.7–32.7 **S1** [turn] No, it's totally not therapy, but, uh, it nudges people to reflect. Think micro-journaling, with maybe prompts like, you know, 'how did you sleep?' or 'what was a win today?' ▷ yields to S3
 28.4–29.3 **S2** [backchannel] Yeah, cool.
 33.2–34.8 **S3** [turn] Ah, like bite-sized reflection. ▷ yields to S1
 34.9–40.7 **S1** [turn] Exactly! And it gently learns, so if, say, someone is struggling with, like, sleep, it shifts more focus there—▷ holds, pauses
 37.3–39.1 **S2** [failed claim] Sorry, do you mean it gives advice?
 41.2–46.1 **S1** [turn] —not exactly advice! More about, like, surfacing patterns or positive reinforcement. ▷ yields to S2
 45.0–46.0 **S3** [backchannel] Right, makes sense.
 46.4–52.8 **S2** [turn] I think—so, what would make somebody want to use this, you know, beyond the first week? I mean, uh, habit apps can be, uh, kind of a tough sell. ▷ yields to S1
 53.2–64.7 **S1** [turn] Yeah, totally, and that's, like, the hardest part. I'm thinking, maybe, subtle rewards or even, um, a community wall with anonymous encouragements? Plus, fewer notifications so it's not overwhelming. ▷ yields to S3
 59.3–60.0 **S2** [backchannel] Mmm, that's—
 64.6–69.2 **S3** [turn] I love the community wall idea, actually. That feels, uh, positive. ▷ yields to S1
 69.7–73.7 **S1** [turn] Thanks! I hope so. And, like, everyone's wall could be slightly different, based on their own journey. ▷ cut off
 72.7–73.6 **S2** [interruption] Can I just—▷ takes floor from S1
 73.8–79.8 **S2** [turn] —so I was thinking, uh, do you imagine this, like, connecting to wearables too? To auto-track mood or sleep or those things? ▷ yields to S1
 79.8–88.6 **S1** [turn] Oh, yeah, I'd love that, honestly. But, um, privacy's tricky, so maybe only if the user really wants to link it. Otherwise, it stays, like, super minimal.
 82.0–83.6 **S2** [laughter] Haha, yeah, no one wants—
 83.4–83.9 **S3** [backchannel] right
 88.6–90.6 **S1** [turn] —no one wants a creepy app, yeah. ▷ yields to S3
 89.3–90.1 **S2** [backchannel] Exactly!
 91.1–97.0 **S3** [turn] So, Anna, if you had, like, one sentence to sum up why someone should download it, what would you say? ▷ yields to S1
 97.0–104.5 **S1** [turn] Um, I guess I'd say, it's a gentle daily nudge for self-compassion—without pressure or judgment. That's, like, the vibe I'm hoping for. ▷ yields to S2
 98.9–99.4 **S3** [backchannel] Nice.
 104.9–106.4 **S2** [turn] Yeah, I really like that. ▷ yields to S3
 106.9–111.6 **S3** [turn] That was super clear. Uh, how did that feel? Want to do it again or get feedback? ▷ yields to S1
 111.5–115.3 **S1** [turn] I'd love feedback, honestly. You guys always spot things I miss. ▷ yields, end
 113.5–115.1 **S2** [failed claim] Totally. Uh, happy to help.

test split · tone: curious and inquisitive · noise: outdoor park, 35.3 dB SNR; chairs scraping · voices: S1 F, S2 F, S3 F
EOT 22 yields / 0 holds · INT 0 interruptions / 6 negatives (5 backchannels, 1 laughter) · 3 failed claims (excluded)

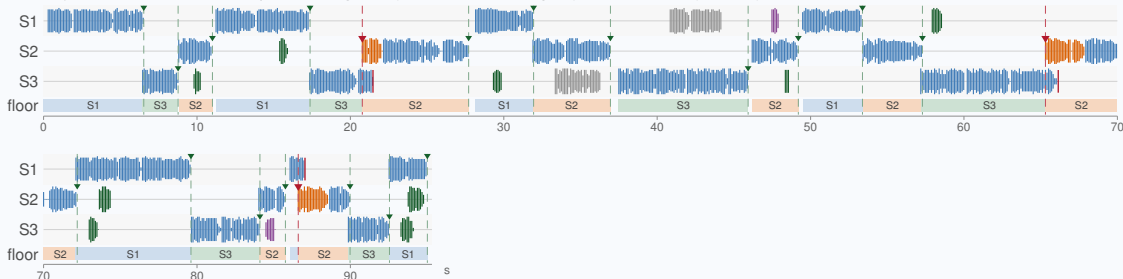


Transcript of test_00186

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–5.6 **S1** [turn] Alright, so, we need to figure out dinner. It's already getting late, so can we just pick something fast? ▷ yields to S2
5.6–9.2 **S2** [turn] Yeah, totally. Are you still craving Thai, or did you change your mind? ▷ yields to S3
9.1–11.3 **S3** [turn] Wait, didn't you both have Thai, like, two nights ago? ▷ yields to S1
10.3–10.8 **S2** [backchannel] Uh-huh
11.3–14.7 **S1** [turn] I did, but I honestly could eat Thai any night, it's not an issue. ▷ yields to S2
14.7–18.9 **S2** [turn] Same, but if you want something else, we can mix it up. What are you feeling, Sam? ▷ yields to S3
18.8–21.8 **S3** [turn] Honestly, I'm kinda pizza-ed out after last weekend. ▷ yields to S1
20.3–21.7 **S1** [backchannel] Oh same, yeah.
21.0–21.6 **S2** [backchannel] Right.
22.4–24.5 **S1** [turn] So, not pizza. What about sushi? ▷ yields to S2
24.5–27.9 **S2** [turn] Ooh, sushi's a good idea. Is Sora still open on Tuesdays?
26.3–27.8 **S3** [failed claim] I think they're closed. Wait—
28.1–31.0 **S2** [turn] —oh wait, you're right, their sign says closed Tuesdays. ▷ yields to S3
31.0–34.2 **S3** [turn] Yeah, definitely. What about that ramen place by the movie theater? ▷ yields to S1
34.2–38.9 **S1** [turn] I haven't tried it yet. Is it actually good? I heard the broth's, like, super salty. ▷ yields to S3
38.8–41.1 **S3** [turn] It's fine. Not amazing, but quick and decent. ▷ yields to S2
41.0–42.5 **S2** [turn] I'm down for ramen. ▷ yields to S1
42.5–49.0 **S1** [turn] Okay, but if we're doing ramen, I need to check they have veggie options. Last time I got stuck with, like, plain noodles. ▷ yields to S2
49.0–51.2 **S2** [turn] Oh true, I forgot. Uh, wanna call them? ▷ yields to S3
51.2–54.5 **S3** [turn] Wait, let me look it up. I think their online menu lists everything.
52.6–54.2 **S2** [failed claim] Thanks, that would actually be—
54.6–57.6 **S3** [turn] —yeah, one sec. Okay, here. They do a spicy miso with tofu. ▷ yields to S1
57.5–59.0 **S1** [turn] Yes! That's great for me. ▷ yields to S3
58.4–58.9 **S2** [backchannel] Mm-hm.
59.3–62.9 **S3** [turn] Cool. So we're set on ramen? Or is anyone still holding out for something else? ▷ yields to S1
60.9–63.3 **S2** [failed claim] I'm good with ramen if everyone else is—
62.9–66.8 **S1** [turn] —yeah, it's fine. We've spent longer talking than we will eating, honestly. ▷ yields to S3
64.7–66.0 **S2** [laughter] Haha, exactly.
67.1–70.6 **S3** [turn] Alright, I'm putting in our order now before someone changes their mind. ▷ yields to S1
70.6–71.3 **S1** [turn] Thank you! ▷ yields to S2
71.3–72.6 **S2** [turn] You're the best, Sam. ▷ yields to S1
72.0–72.7 **S3** [backchannel] Yeah, yeah.
72.9–75.1 **S1** [turn] Don't forget extra ginger, please. ▷ yields, end

test split · tone: professional and business-like · noise: quiet apartment, 29.6 dB SNR; laughter offmic, alarm beep · voices: S1 M, S2 F, S3 F
EOT 18 yields / 0 holds · INT 3 interruptions / 11 negatives (9 backchannels, 2 laughter) · 2 failed claims (excluded)



Transcript of test_00188

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.5 **S1** [turn] Alright, so I guess we're starting! Uh, did everyone finish The Midnight Library? I actually powered through it last night. ▷ yields to S3
6.5–8.8 **S3** [turn] Yeah, I finished it two days ago, I just— ▷ yields to S2
8.8–11.0 **S2** [turn] Of course. I read it in one sitting. ▷ yields to S1
9.8–10.3 **S3** [backchannel] wow
11.2–17.4 **S1** [turn] That's impressive. I had to stop a few times. You know, those decisions she had to make, they really made me stop and think. ▷ yields to S3
15.4–15.9 **S2** [backchannel] Mm-hm.
17.3–21.5 **S3** [turn] Same here, I got really emotional about her regrets. The way the library worked was so— ▷ cut off
20.8–22.1 **S2** [interruption] It was formulaic. ▷ takes floor from S3

22.2–27.7 **S2** [turn] Honestly, the structure didn't surprise me after the first few chapters. It telegraphed most twists. > yields to S1

28.2–31.9 **S1** [turn] Yeah, that's fair. I kinda liked knowing where it was going, though. Like, it felt safe? > yields to S2

29.3–29.9 **S3** [backchannel] mm-hm

31.9–34.0 **S2** [turn] I need a bit more challenge, personally.

33.4–36.4 **S3** [failed claim] But the comfort factor was kind of the point, wasn't it?

34.1–37.0 **S2** [turn] —so I'll admit, it was an easy read for a rough week. > yields to S3

37.5–41.9 **S3** [turn] Right, exactly. I found those little moments—like with her brother—super comforting.

40.9–44.2 **S1** [failed claim] Oh, that scene got me! When they played music together in that one life?

42.0–45.9 **S3** [turn] —yeah, and she finally found that sense of belonging? That really moved me. > yields to S2

46.2–49.2 **S2** [turn] I didn't find it as affecting but I get why people do. > yields to S1

47.5–47.9 **S1** [laughter] haha

48.4–48.6 **S3** [backchannel] mm

49.5–53.4 **S1** [turn] Did either of you imagine what you'd pick if you landed in a library like that? > yields to S2

53.4–57.3 **S2** [turn] Of course. I'd grab the book where I chose engineering over finance. > yields to S3

57.2–60.1 **S3** [turn] That's so interesting. I kept thinking about that too—

57.9–58.6 **S1** [backchannel] Yeah, same.

60.2–66.1 **S3** [turn] maybe something where I moved to Spain after college, or tried something totally out of my comfort zone. It's kind of freeing to think about. > cut off

65.3–67.9 **S2** [interruption] Except, it's also a little—what, paralyzing? > takes floor from S3

67.9–72.2 **S2** [turn] Thinking about every possible outcome. I get why she felt overwhelmed. > yields to S1

72.1–74.9 **S1** [turn] Yeah, too many doors. Sometimes it's better not to know, right?

73.0–73.6 **S3** [backchannel] Right.

73.6–74.4 **S2** [backchannel] Agreed.

74.9–79.6 **S1** [turn] Did the ending work for you both? I thought it was a little, uh, easy maybe, but satisfying. > yields to S3

79.6–84.1 **S3** [turn] Yeah, I felt good closing the book. It wrapped up but still left things open enough. > yields to S2

84.0–85.8 **S2** [turn] It was serviceable. Tidy. > yields to S1

84.5–85.1 **S3** [laughter] haha

86.1–87.0 **S1** [turn] Alright, so next month— > cut off

86.6–88.6 **S2** [interruption] Let's pick something more challenging. > takes floor from S1

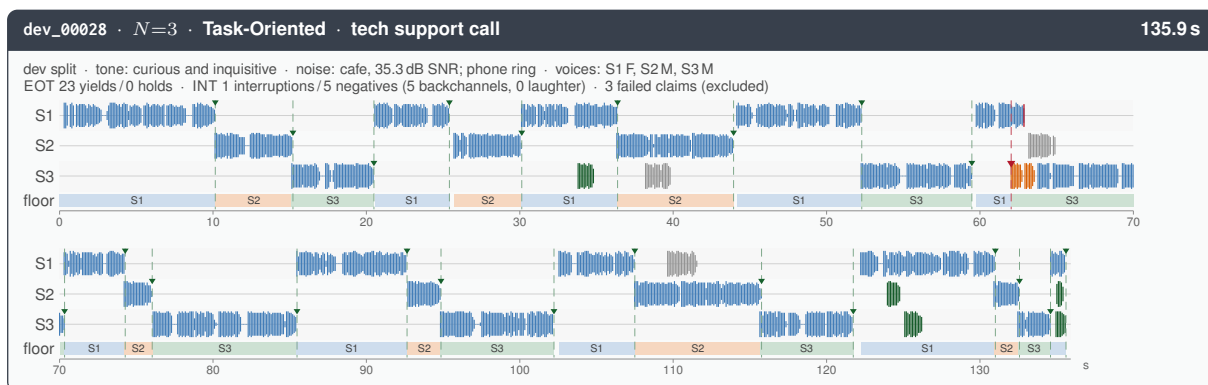
88.7–90.0 **S2** [turn] Maybe nonfiction? > yields to S3

89.9–92.5 **S3** [turn] I'm open to that. As long as it isn't finance. > yields to S1

92.5–95.0 **S1** [turn] Haha, deal. No finance. I'll send around a survey? > yields, end

93.3–94.1 **S3** [backchannel] Perfect.

93.8–94.8 **S2** [backchannel] Sounds good.



Transcript of dev_00028

start-end (s) speaker [ground-truth label] text > outcome

0.3–10.2 **S1** [turn] Hi! Uh, thanks for joining the call, both of you. So my laptop's wifi keeps dropping, especially when I try to connect to the office VPN. Can we figure out what's going on? > yields to S2

10.1–15.2 **S2** [turn] Yeah, totally! Let's start with the basics. You said the wifi drops mostly when you use the VPN, right? > yields to S3

15.1–20.5 **S3** [turn] That's already a useful clue. Uh, can you remind me what OS version you have running? > yields to S1

20.6–25.4 **S1** [turn] It's Windows eleven, um, with the latest updates, I think. I just did one last week. > yields to S2

25.7–30.1 **S2** [turn] Perfect, thanks. And have you noticed if it only happens on your home network or at the office too? > yields to S1

30.1–36.4 **S1** [turn] Oh, interesting, um — it's mostly at home, actually. I don't think it's ever dropped at the office, now that you mention it. > yields to S2

33.8–34.9 **S3** [backchannel] That's helpful.

36.3–41.2 **S2** [turn] Right, so maybe it's something about your home router or, uh, maybe the VPN is more sensitive there.

38.2–39.9 **S3** [failed claim] Is it okay if I—

41.3–43.9 **S2** [turn] —ask if you ever get an error message when it disconnects? > yields to S1

44.1–52.3 **S1** [turn] Sometimes, yeah! There's this little pop-up that says something like, uh, 'Network connection lost, reconnecting,' and then it'll try again by itself. > yields to S3

52.3–59.5 **S3** [turn] Hmm, that could be the wifi adapter resetting. Do other devices lose connection at the same time, or just your laptop? > yields to S1

59.8–62.9 **S1** [turn] Just my laptop, I think. My phone and tablet stay online. It's super annoying. > cut off

62.0–63.6 **S3** [interruption] Can I jump in a sec? > takes floor from S2

63.2–65.0 **S2** [failed claim] That narrows it down a lot. Uh—

63.8–70.3 **S3** [turn] I'd like to check device manager logs—have you ever looked for errors in device manager after a disconnect? > yields to S1

70.3–74.3 **S1** [turn] Uh, I haven't, no. I didn't even know you could do that. Where would I find those? > yields to S2

74.2–76.0 **S2** [turn] Yeah, it's under Network Adapters— > yields to S3

76.1–85.5 **S3** [turn] Exactly, you just right-click your wifi adapter, choose Properties, then check the Events tab. If you see any red Xs or warnings, let us know. > yields to S1

85.5–92.7 **S1** [turn] Okay, hold on, let me try—um, Properties... oh! There's a warning here, it says something about 'Device reset due to network timeout.' > yields to S2

92.7–94.9 **S2** [turn] Oh wow, that could really be a driver issue. > yields to S3

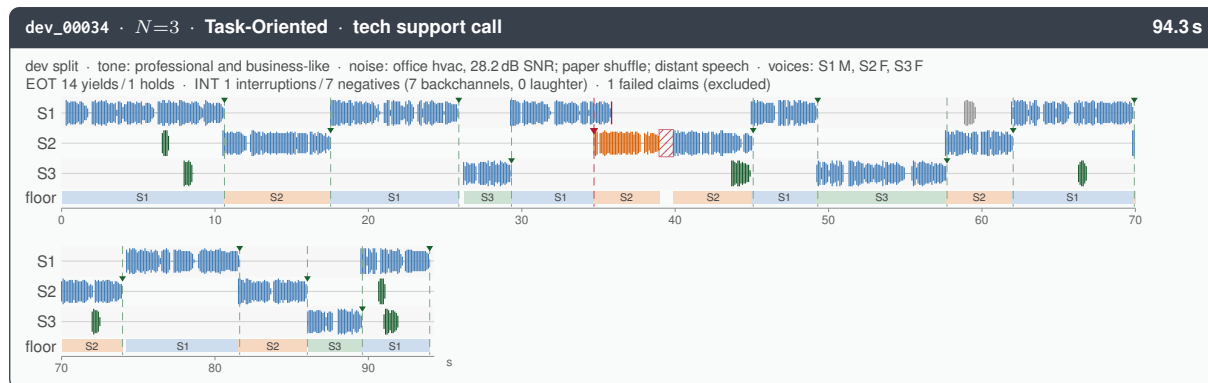
94.8–102.2 **S3** [turn] Or maybe some router compatibility thing. Do you know if your router supports both two point four and five gigahertz bands? > yields to S1

102.6–107.5 **S1** [turn] Um, I think so? It has two wifi names, like one says '2G' and the other's '5G'. > yields to S2

107.5–110.5 **S2** [turn] Right. Have you tried switching between them, just to see if it helps?

109.6–111.6 **S1** [failed claim] I haven't, but I can try. Should I—

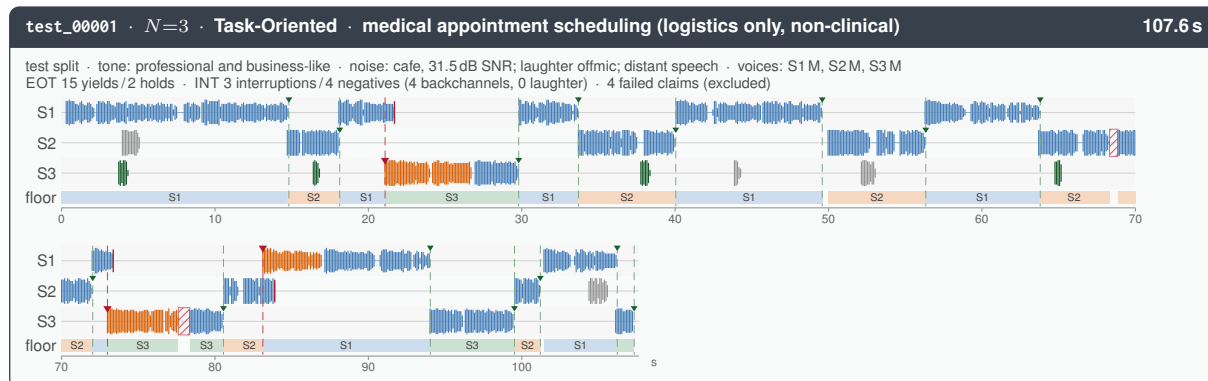
110.5–115.8 **S2** [turn] —yeah, maybe just connect to the two point four one for a while, especially on VPN, and see if it's more stable. ▷ yields to S3
 115.7–121.7 **S3** [turn] And update your network driver if you haven't. Sometimes Windows updates leave that behind. ▷ yields to S1
 122.2–127.5 **S1** [turn] Okay, I'll try those! Uh, and if it still drops, should I, like, screenshot the warning and send it to you?
 124.0–124.8 **S2** [backchannel] Yes, please.
 125.1–126.2 **S3** [backchannel] That's perfect.
 127.6–131.0 **S1** [turn] Cool, I guess I'll test it tonight. Thanks so much, both of you! ▷ yields to S2
 130.9–132.6 **S2** [turn] No problem at all, happy to help! ▷ yields to S3
 132.5–134.6 **S3** [turn] Good luck, let us know how it goes. ▷ yields to S1
 134.6–135.6 **S1** [turn] Will do! Bye! ▷ yields, end
 135.0–135.6 **S3** [backchannel] Bye!
 135.0–135.5 **S2** [backchannel] Bye!



Transcript of dev_00034

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–10.6 **S1** [turn] Thanks for joining, both of you. We've got the server outage ticket for finance, so let's just get straight to triage. I'll start with what I've seen—database connectivity errors, starting just before nine. ▷ yields to S2
 6.6–7.0 **S2** [backchannel] Right.
 8.0–8.6 **S3** [backchannel] Uh-huh.
 10.5–17.5 **S2** [turn] But are you sure it's the database? I got a couple pings from users saying they couldn't even load the web portal at all, not just failed lookups. ▷ yields to S1
 17.5–25.9 **S1** [turn] Yeah, but the application logs point straight to SQL timeouts. I checked the web servers—they're healthy, except for those failures when they hit finance-db. ▷ yields to S3
 26.2–29.3 **S3** [turn] Okay, um, did anyone check for a network segment outage? ▷ yields to S1
 29.3–35.9 **S1** [turn] Already on it, nothing in the monitoring dashboard. All pings to the DB server respond fine from our test boxes. ▷ cut off
 34.7–38.9 **S2** [interruption] Wait, did you actually run a trace from inside the app subnet or just from infra? ▷ takes floor from S1 ▷ holds, pauses
 39.9–45.1 **S2** [turn] Because sometimes the overlay network does weird things with routing between those zones, seriously. ▷ yields to S1
 43.7–44.9 **S3** [backchannel] Oh, yeah, totally.
 45.0–49.3 **S1** [turn] I just did infra to DB. That's a good point—I'll try from an app host now. ▷ yields to S3
 49.2–57.7 **S3** [turn] So, um, while he checks that, I noticed the finance-db CPU spiked right before the failures. Like, over ninety percent for a few minutes. ▷ yields to S2
 57.6–59.8 **S2** [turn] Interesting. Anything in the SQL logs?
 58.9–59.7 **S1** [failed claim] Hang on—
 60.0–62.0 **S2** [turn] —or did it just flat-out stop logging? ▷ yields to S1
 61.9–69.9 **S1** [turn] —it froze up around 8:57, tons of slow queries suddenly. Looks like a reporting job or maybe someone ran a big ad hoc statement? ▷ yields to S2
 66.3–66.8 **S3** [backchannel] Yeah.
 69.8–74.0 **S2** [turn] That matches what I'm hearing from finance. They couldn't save anything after nine. ▷ yields to S1
 72.0–72.5 **S3** [backchannel] Wow.
 74.2–81.6 **S1** [turn] Okay, network from app subnet looks fine, so yeah, this is on the DB side. Should we check if someone's still running that reporting job? ▷ yields to S2
 81.5–86.0 **S2** [turn] I'll call finance and see who kicked it off, and I'll ask them to cancel if it's still running. ▷ yields to S3
 86.0–89.6 **S3** [turn] Do you want me to monitor CPU? I can alert if it spikes again. ▷ yields to S1
 89.5–94.0 **S1** [turn] Perfect, thanks both. Let's touch base in ten? Ping if anything changes sooner. ▷ yields, end
 90.7–91.2 **S2** [backchannel] Yep.
 91.0–92.0 **S3** [backchannel] Sounds good!



Transcript of test_00001

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.1 **S1** [turn] Alright, thanks for joining, guys. The main thing today is to coordinate Dr. Lee's schedule for next week since the clinic hours got shifted. I'll pull up the calendar—

3.7–4.4 **S3** [backchannel] Yeah, cool.

4.0–5.1 **S2** [failed claim] I'm ready when you are.

9.2–14.8 **S1** [turn] Okay, so—all of Dr. Lee's morning appointments have to be moved to afternoons on Tuesday and Thursday. ▷ yields to S2

14.7–18.1 **S2** [turn] Just to clarify, that's both new and follow-up appointments, right? ▷ yields to S1

16.4–16.8 **S3** [backchannel] Mm-hm.

18.1–21.7 **S1** [turn] Yes, exactly. The whole block from nine to noon, just shift it after lunch. ▷ cut off

21.1–26.8 **S3** [interruption] Wait, what about the people already double-booked on Tuesday? Gail Rogers and Samir, remember those two at one? ▷ takes floor from S1

27.0–29.8 **S3** [turn] I think if we move them all, Dr. Lee won't have enough time. ▷ yields to S1

29.8–33.7 **S1** [turn] That's true—good catch. Maybe we stagger the follow-ups later in the week then. ▷ yields to S2

33.7–40.0 **S2** [turn] Would it work to put the lower-urgency follow-up patients on Friday afternoon? Dr. Lee's schedule looks clearer then. ▷ yields to S1

37.8–38.4 **S3** [backchannel] Right, yeah.

40.1–45.6 **S1** [turn] I agree, Friday's fine. I'll make a note. Uh, for new patients, should we try for Wednesday mornings instead?

43.9–44.3 **S3** [failed claim] But—

45.7–49.6 **S1** [turn] —I mean, unless that's going to create a bottleneck with Dr. Lee's procedure blocks? ▷ yields to S2

50.0–56.3 **S2** [turn] No, I think Wednesday's open until eleven, actually. I checked that before. But we should call to confirm with him. ▷ yields to S1

52.1–53.1 **S3** [failed claim] Yeah, just to be safe.

56.3–63.8 **S1** [turn] Alright, I'll send Dr. Lee a note right after this. For patients who already confirmed for Tuesday, I guess we'll start calling them this afternoon? ▷ yields to S2

63.7–68.3 **S2** [turn] Yes, and we should give at least twenty-four hours' notice. Some of them come in from out of town. ▷ holds, pauses

64.8–65.2 **S3** [backchannel] Uh-huh.

68.9–72.0 **S2** [turn] Oh, do we have the interpreter scheduled for Samir's appointment? ▷ yields to S1

72.0–73.4 **S1** [turn] I saw a note about that— ▷ cut off

73.0–77.6 **S3** [interruption] Hang on, I think it's only for the original morning slot. If we move it, we'll have to rebook. ▷ takes floor from S1 ▷ holds, pauses

78.4–80.6 **S3** [turn] I can call scheduling right after we finish here. ▷ yields to S2

80.6–83.9 **S2** [turn] Perfect, thanks. Let's also double-check that all patient contact numbers are up to date— ▷ cut off

83.1–87.0 **S1** [interruption] —yeah, I'll cross-reference the EMR and send you both the list this afternoon. ▷ takes floor from S2

87.2–94.0 **S1** [turn] Oh! One more thing, there's a new referral that just came in for Friday. Should I try squeezing them in as a last slot, or wait for next week? ▷ yields to S3

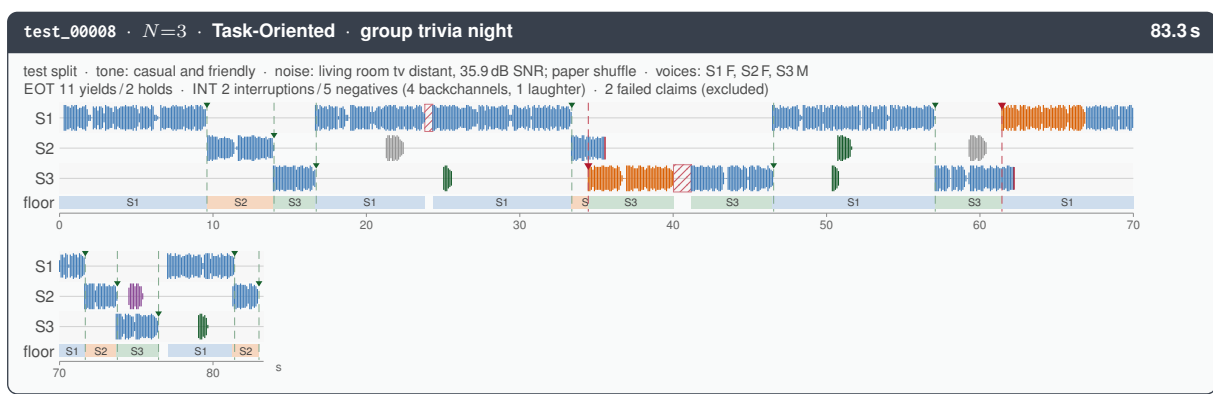
94.0–99.5 **S3** [turn] Let's hold off till next week if possible. Dr. Lee prefers not to rush new patients on Fridays. ▷ yields to S2

99.5–101.2 **S2** [turn] Agreed. Next week is better. ▷ yields to S1

101.4–106.2 **S1** [turn] Alright, I think that covers everything. Thanks, guys. I'll confirm the changes in email. ▷ yields to S3

104.4–105.6 **S2** [failed claim] Sounds good, thank you.

106.1–107.3 **S3** [turn] Alright, thanks! See you later. ▷ yields, end



Transcript of test_00008

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.6 **S1** [turn] Okay, so question seven: what was the first animated film to be nominated for a Best Picture Oscar? I'm pretty sure I know, but does anyone else have a guess? ▷ yields to S2

9.7–14.0 **S2** [turn] I wanna say Beauty and the Beast. That rings a bell, but I'm not 100 percent sure. ▷ yields to S3

13.9–16.7 **S3** [turn] Yeah, I think it's that one, um, from '91, right? ▷ yields to S1

16.7–23.8 **S1** [turn] Yep, that's what I had in my head too. I remember it was kind of a big deal, because it was up against Silence of the Lambs, I think? ▷ holds, pauses

21.3–22.5 **S2** [failed claim] Right, that sounds—

24.3–27.3 **S1** [turn] —like the year. Okay, going with Beauty and the Beast.

25.0–25.6 **S3** [backchannel] mm-hm

27.4–33.4 **S1** [turn] Alright, next one is a geography one. Oh, this is tricky—what's the only country that's also a continent? ▷ yields to S2

33.4–35.6 **S2** [turn] Australia. That's gotta be Australia, right? ▷ cut off

34.5–40.0 **S3** [interruption] Wait, so is Antarctica counted for that? Or is it just Australia because it's a country and a continent? ▷ takes floor from S2 ▷ holds, pauses

41.2–46.6 **S3** [turn] I mean, Antarctica isn't a country, right? It's just a continent, so—I think Australia makes sense. ▷ yields to S1

46.5–52.3 **S1** [turn] Yeah, it's gotta be Australia. Antarctica's just, like, a landmass without a government or anything.

50.4–50.8 **S3** [backchannel] uh-huh

50.7–51.7 **S2** [backchannel] Exactly.

52.4–57.1 **S1** [turn] Alright, locking in Australia. Oh, next one's for you, Jamie! Who painted Starry Night? ▷ yields to S3

57.1–62.2 **S3** [turn] Uh, that's Van Gogh. Vincent van Gogh. I always think of, um, the swirly sky and the, like, yellow moon. ▷ cut off

59.3–60.4 **S2** [failed claim] Yeah, you got it.

61.4–66.9 **S1** [interruption] Nice! Okay, next—oh, wait, can I skip ahead? The music round's my favorite. Can we do number twelve? ▷ takes floor from S3

67.0–71.7 **S1** [turn] Which Beatles album had the song "Yesterday"? I always forget—was it Help! or Rubber Soul? ▷ yields to S2

71.6–73.8 **S2** [turn] Help!. I'm like, ninety-nine percent sure— ▷ yields to S3

73.7–76.5 **S3** [turn] Yeah, Help!. Definitely. My dad had that on vinyl. ▷ yields to S1

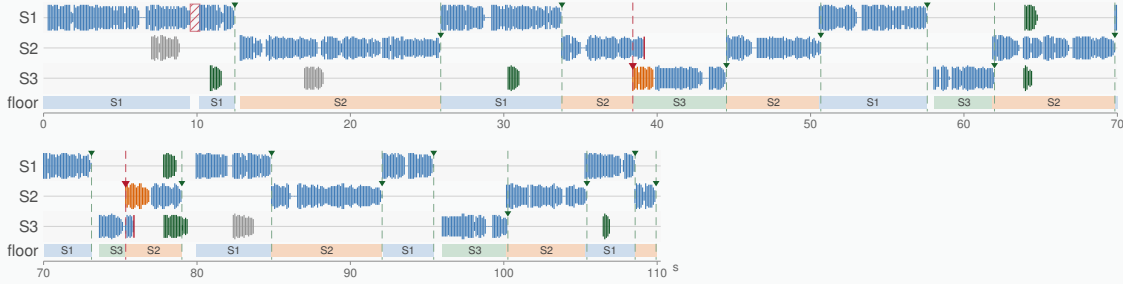
74.5–75.4 **S2** [laughter] haha classic

77.0–81.4 **S1** [turn] Alright, Help!. You're all sure? Good, marking it. We're, like, crushing this round, you guys. ▷ yields to S2

79.1–79.7 **S3** [backchannel] woo

81.3–83.0 **S2** [turn] Yeah, we're on a roll. Next one? ▷ yields, end

test split · tone: professional and business-like · noise: outdoor park, 27.6 dB SNR; school bell · voices: S1 F, S2 F, S3 F
EOT 17 yields / 1 holds · INT 2 interruptions / 7 negatives (7 backchannels, 0 laughter) · 3 failed claims (excluded)



Transcript of test_00012

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.6 **S1** [turn] Alright, so just to kick us off, our key deliverable is a draft itinerary for the Tokyo trip by end of day. I've pulled the client wishlist and hotel details— ▷ holds, pauses

7.0–8.9 **S2** [failed claim] I have the flight options ready—

10.2–12.5 **S1** [turn] —so we can slot those in first. Sound good? ▷ yields to S2

10.9–11.7 **S3** [backchannel] Absolutely!

12.8–25.9 **S2** [turn] Let's anchor on the flight dates. The client prefers departing the 14th and returning the 21st. I've got ANA 11:00 a.m. direct for the outbound, JAL 5:25 p.m. on the return—does everyone see an issue? ▷ yields to S1

17.0–18.2 **S3** [failed claim] Nope, those look great.

25.9–33.8 **S1** [turn] Perfect, that gives us almost a full week to play with. Hotel is confirmed at the Shinjuku Granbell, two queens, breakfast included. ▷ yields to S2

30.3–31.1 **S3** [backchannel] Ooh, fancy!

33.8–39.2 **S2** [turn] Moving on—activities. The client's top picks were Tsukiji Market, Studio Ghibli, and a sumo match if available. ▷ cut off

38.4–39.8 **S3** [interruption] I checked the sumo schedule— ▷ takes floor from S2

39.9–44.5 **S3** [turn] —there's actually a tournament on the 17th at Ryogoku. Should I pencil that in? ▷ yields to S2

44.5–50.7 **S2** [turn] Yes, please block that morning. Then we can do the Ghibli Museum in the afternoon, assuming we can snag tickets. ▷ yields to S1

50.6–57.6 **S1** [turn] I can monitor the release times—they go fast. I'll try for the earliest slot on the 17th or 18th if the sumo sells out. ▷ yields to S3

58.0–62.0 **S3** [turn] Thanks! And, um, for Tsukiji, is the outer market still open to tours? ▷ yields to S2

61.9–65.1 **S2** [turn] Yes, but not the inner auctions. We can book a food tour—

63.9–64.5 **S3** [backchannel] Yessss.

64.0–64.8 **S1** [backchannel] Oh, love that.

65.2–69.8 **S2** [turn] —for the 15th, morning slot. Then maybe Akihabara electronics shopping after? ▷ yields to S1

69.8–73.1 **S1** [turn] That works, and leaves the evening open for something spontaneous. ▷ yields to S3

73.6–75.9 **S3** [turn] I was thinking robot restaurant! Kitsch but— ▷ cut off

75.4–76.9 **S2** [interruption] It's closed, actually. ▷ takes floor from S3

77.1–79.0 **S2** [turn] Pandemic casualty, sorry. ▷ yields to S1

77.8–78.7 **S1** [backchannel] Oh no, really?

77.9–79.4 **S3** [backchannel] RIP robot dancers.

79.9–84.9 **S1** [turn] Ok, we'll find another dinner show. For the last day, thoughts on a day trip to Hakone? ▷ yields to S2

82.4–83.7 **S3** [failed claim] Hot springs, yes please!

84.9–92.1 **S2** [turn] Logistics could be tight. Bullet train is 36 minutes, but we'll need to book lockers for luggage, unless we leave them at the hotel. ▷ yields to S1

92.1–95.4 **S1** [turn] I'll look into luggage forwarding. Otherwise, yeah, lockers. ▷ yields to S3

96.0–100.3 **S3** [turn] Can we squeeze in the teamLab digital art thing somewhere? Maybe the 16th? ▷ yields to S2

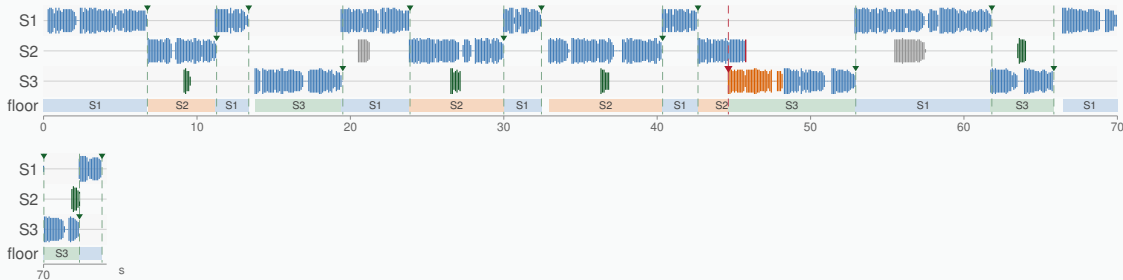
100.2–105.4 **S2** [turn] I think that's doable if we skip the Meiji Shrine on the 16th afternoon and go straight to teamLab. ▷ yields to S1

105.3–108.6 **S1** [turn] Sold. I'll draft this up and circulate by three. Sound good? ▷ yields to S2

106.5–106.9 **S3** [backchannel] Yep!

108.5–109.9 **S2** [turn] Perfect, thanks all. ▷ yields, end

test split · tone: casual and friendly · noise: cafe, 25.8 dB SNR; door knock · voices: S1 F, S2 M, S3 F
EOT 15 yields / 0 holds · INT 1 interruptions / 5 negatives (5 backchannels, 0 laughter) · 2 failed claims (excluded)



Transcript of test_00086

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.8 **S1** [turn] Alright, so thanks for hopping on! First thing—are we all set for Saturday's park clean-up, or is anything up in the air? ▷ yields to S2

6.8–11.3 **S2** [turn] I think we're mostly good, but um, transportation for the supplies isn't fully sorted yet. ▷ yields to S1

9.2–9.6 **S3** [backchannel] right

11.2–13.4 **S1** [turn] Oh, do we still need a truck for the mulch bags? ▷ yields to S3

13.8–19.5 **S3** [turn] Yeah, and plus the rakes and gloves are still at the rec center, so someone’s gotta pick those up Friday. > yields to S1

19.4–21.9 **S1** [turn] I can do the rec center run, that’s no problem.

20.5–21.3 **S2** [failed claim] Uh, but—

22.0–23.9 **S1** [turn] —unless someone else already volunteered? > yields to S2

23.8–30.0 **S2** [turn] No, I was just gonna say the rec center closes early on Fridays, so just, uh, double check the pick-up hours. > yields to S1

26.6–27.3 **S3** [backchannel] good point

30.0–32.5 **S1** [turn] Awesome heads up, thanks! I’ll call them tomorrow then. > yields to S2

32.9–40.4 **S2** [turn] Now, on transportation—Jackie offered her SUV for mulch, but there’s no trailer. Um, I could ask my brother if he’s free with his truck though. > yields to S1

36.3–36.9 **S3** [backchannel] oh nice

40.3–42.7 **S1** [turn] Would he be cool with that? Or does he charge, haha? > yields to S2

42.7–45.8 **S2** [turn] Haha, no charge! He just likes getting out, so I think he’ll help if I ask. > cut off

44.6–48.2 **S3** [interruption] I actually think we need to finalize the volunteer roles, too— > takes floor from S2

48.3–52.9 **S3** [turn] —because last time some folks just sort of wandered, you know? Can we assign groups this time? > yields to S1

52.9–58.4 **S1** [turn] Oh, for sure. I was thinking like, team captains for each area—mulch, trash, and gardens. Sound good?

55.5–57.5 **S2** [failed claim] Yeah, uh, I’d be okay leading trash—

58.4–61.8 **S1** [turn] —legend! I’ll do gardens then, and speaker_3, you good with mulch? > yields to S3

61.7–65.9 **S3** [turn] Yep, I can wrangle the mulch crew. I’ll bring the sign-up sheets and markers. > yields to S1

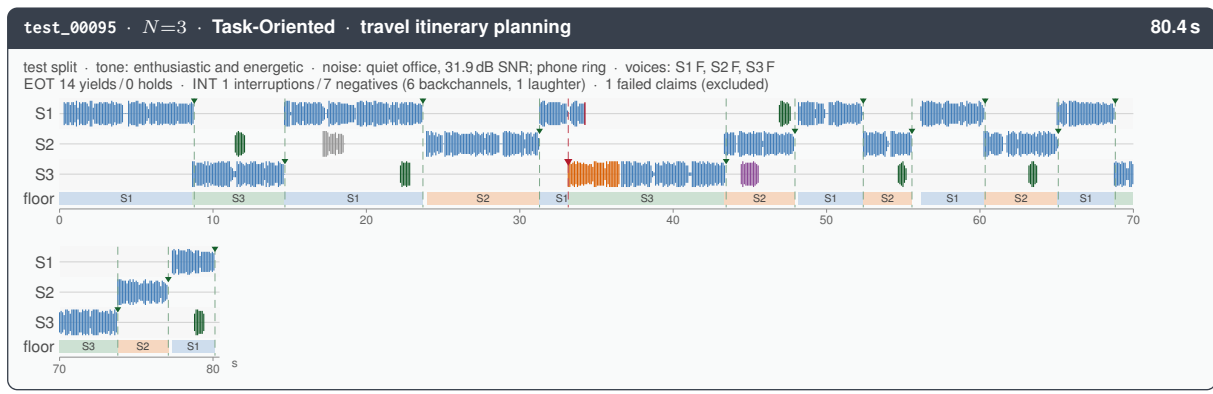
63.5–64.1 **S2** [backchannel] perfect

66.4–70.0 **S1** [turn] Alright, I’ll draft up a summary and send it out tonight. Anything else? > yields to S3

70.1–72.3 **S3** [turn] Nope, I think that covers it. Thanks, all! > yields to S1

71.8–72.4 **S2** [backchannel] thanks!

72.3–73.8 **S1** [turn] You guys rock, see you Saturday! > yields, end



Transcript of test_00095

start–end (s) speaker [ground-truth label] text > outcome

0.3–8.8 **S1** [turn] Okay, so, big picture, we’ve got three full days in Barcelona—what should we lock in first, like do you both want to put the Sagrada Familia at the top? > yields to S3

8.7–14.7 **S3** [turn] Absolutely, that has to be on day one, I think. It’s always busy so maybe we book that for the morning slot? > yields to S1

11.5–12.1 **S2** [backchannel] yeah

14.7–19.3 **S1** [turn] Love it. Early tickets, get it done before the crowds—then brunch somewhere nearby maybe?

17.2–18.6 **S2** [failed claim] Oh, speaking of food—

19.4–23.7 **S1** [turn] —I found this café called Brunch & Cake, looks so cute, and it’s right down the street. > yields to S2

22.2–22.9 **S3** [backchannel] oh wow

23.9–31.3 **S2** [turn] That sounds perfect! Okay, Sagrada Familia in the morning, Brunch & Cake for late breakfast. What about the afternoon—is Parc Güell good for then? > yields to S1

31.3–34.2 **S1** [turn] Ooh yes, and Gaudi overload, love it. Should we walk there or— > cut off

33.2–36.5 **S3** [interruption] Wait, isn’t that a really steep hike if we walk from Sagrada? > takes floor from S1

36.7–43.5 **S3** [turn] I mean, I’m up for it, but just checking—do we want to save energy and grab a bus? I don’t want anyone to get too exhausted. > yields to S2

43.3–47.9 **S2** [turn] Totally fair! Uh, maybe we bus it up and then walk back down, so we get the views without dying? > yields to S1

44.4–45.6 **S3** [laughter] haha perfect

46.9–47.7 **S1** [backchannel] exactly

48.2–52.4 **S1** [turn] Okay, what about dinner the first night? Tapas crawl in El Born, or something different? > yields to S2

52.4–55.6 **S2** [turn] Tapas crawl, one hundred percent. And maybe churros for dessert? > yields to S1

54.7–55.2 **S3** [backchannel] yes

56.1–60.3 **S1** [turn] Sold! Okay, day two, do we want to do the beach or the Picasso Museum first? > yields to S2

60.3–65.1 **S2** [turn] I vote beach in the morning—less crowded, and it leaves the museum for when it’s hot in the afternoon. > yields to S1

63.2–63.8 **S3** [backchannel] mm-hm

65.0–68.8 **S1** [turn] That’s smart! Oh, do you both want to try paddleboarding, or just swim and chill? > yields to S3

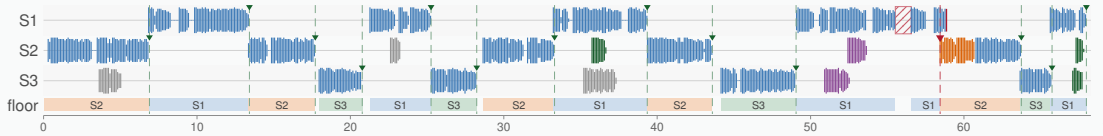
68.8–73.8 **S3** [turn] Can we do both? Like swim first and then maybe I’ll work up the nerve for paddleboarding after? > yields to S2

73.8–77.1 **S2** [turn] haha yes, I am right there with you, let’s see how we feel after coffee! > yields to S1

77.4–80.1 **S1** [turn] yes! Okay, this is already shaping up to be my dream trip. > yields, end

78.8–79.4 **S3** [backchannel] same!

test split · tone: casual and friendly · noise: outdoor park, 29.0 dB SNR; keyboard clatter · voices: S1 F, S2 F, S3 F
EOT 13 yields / 1 holds · INT 1 interruptions / 5 negatives (3 backchannels, 2 laughter) · 3 failed claims (excluded)

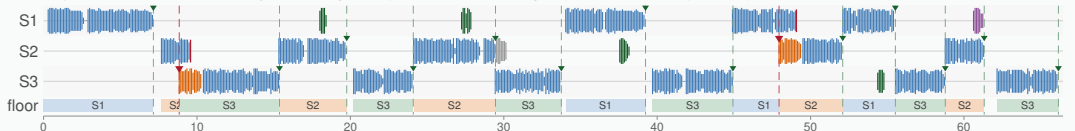


Transcript of test_00113

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.9 **S2** [turn] Okay, so are we doing Codenames or Ticket to Ride? I brought both, but Codenames is, like, easier if we're still eating. ▷ yields to S1
3.6–5.1 **S3** [failed claim] Ooh, I love Codenames.
6.8–13.4 **S1** [turn] I'd vote Codenames too. Less setup, and I feel like we never finish Ticket to Ride before someone has to leave. ▷ yields to S2
13.4–17.7 **S2** [turn] True, that's a good point. Alright, Codenames it is then! Who wants to be spymaster? ▷ yields to S3
18.0–20.8 **S3** [turn] Not it! I always panic when I have to come up with clues. ▷ yields to S1
21.3–23.8 **S1** [turn] I don't mind starting as spymaster, unless you—
22.6–23.3 **S2** [failed claim] Wait, wait—
23.9–25.3 **S1** [turn] —really want to, obviously. ▷ yields to S3
25.3–28.2 **S3** [turn] No, you go ahead! You're way better at those connections than me. ▷ yields to S2
28.7–33.3 **S2** [turn] Yeah, seriously, you have those galaxy brain clues. Okay, so we'll split into teams— ▷ yields to S1
33.2–39.4 **S1** [turn] I think with three, it makes sense for me to be spymaster for both teams, right? It gets a little tricky, but— ▷ yields to S2
35.2–37.3 **S3** [failed claim] That's the way we did it last time, yeah.
35.8–36.7 **S2** [backchannel] Yeah, exactly.
39.4–43.6 **S2** [turn] But you didn't want to try that variant where you swap spymasters each round, did you? ▷ yields to S3
44.2–49.1 **S3** [turn] Oh, that could be fun! But maybe we should wait until after this round so I can remember the rules again. ▷ yields to S1
49.1–55.5 **S1** [turn] Totally reasonable. Okay, give me like thirty seconds to lay out the cards. Don't look at this side, please. ▷ holds, pauses
50.9–52.5 **S3** [laughter] Haha, yeah, eyes closed.
52.4–53.7 **S2** [laughter] Haha, promise I won't peek.
56.6–58.9 **S1** [turn] Alright, all set. Who wants to guess first, or should I just pick randomly? ▷ cut off
58.4–60.7 **S2** [interruption] Wait, can I go first? I'm feeling lucky tonight— ▷ takes floor from S1
60.8–63.7 **S2** [turn] —I lost at literally every other game, so I need a win. ▷ yields to S3
63.7–65.7 **S3** [turn] Go for it! I'll just cheer for both sides. ▷ yields to S1
65.6–68.0 **S1** [turn] Okay then, first clue coming up. Ready? ▷ yields, end
67.1–67.8 **S3** [backchannel] Ready.
67.3–67.8 **S2** [backchannel] Ready!

test split · tone: enthusiastic and energetic · noise: living room tv distant, 25.8 dB SNR; paper shuffle, laughter offmic · voices: S1 M, S2 M, S3 F
EOT 13 yields / 0 holds · INT 2 interruptions / 5 negatives (4 backchannels, 1 laughter) · 1 failed claims (excluded)

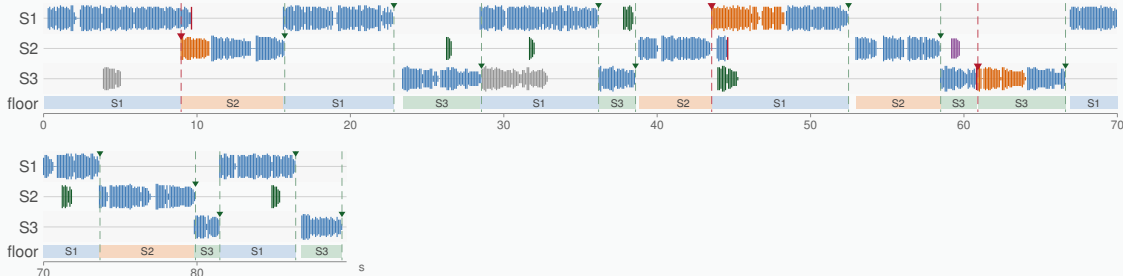


Transcript of test_00130

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–7.2 **S1** [turn] Alright, let's tackle the bio review sheet. Anyone want to start with enzymes or mitosis? I'm good with either, honestly. ▷ yields to S2
7.7–9.6 **S2** [turn] Um, maybe enzymes? I always mix up the— ▷ cut off
8.8–10.3 **S3** [interruption] Active site stuff, right? ▷ takes floor from S2
10.4–15.4 **S3** [turn] Because it's always like, the enzyme fits with the substrate, lock and key, but then there's induced fit too. ▷ yields to S2
15.3–19.8 **S2** [turn] Right, yeah. That's what trips me up. Like, are we supposed to know the difference for this test or— ▷ yields to S3
18.0–18.5 **S1** [backchannel] Uh-huh
20.2–24.1 **S3** [turn] Yeah, she said so. Remember last week? There was that whole slide comparing them. ▷ yields to S2
24.1–29.5 **S2** [turn] Oh, okay. So lock and key is more rigid, induced fit changes shape around the substrate. Got it... I think. ▷ yields to S3
27.2–27.9 **S1** [backchannel] Exactly
29.4–33.8 **S3** [turn] And don't forget inhibitors! Competitive versus noncompetitive. That's always on there. ▷ yields to S1
29.5–30.2 **S2** [failed claim] Wait—
34.0–39.2 **S1** [turn] Yup, competitive goes right at the active site, noncompetitive changes the shape, basically. ▷ yields to S3
37.6–38.2 **S2** [backchannel] Right
39.7–44.9 **S3** [turn] Okay, should we move on to mitosis? That diagram she keeps putting on the quizzes always messes me up. ▷ yields to S1
44.9–49.1 **S1** [turn] Yeah, I can run through it if you want. So, it's PMAT: prophase, metaphase, anaphase, telophase. ▷ cut off
47.9–49.4 **S2** [interruption] Wait, so in metaphase— ▷ takes floor from S1
49.5–52.1 **S2** [turn] That's where the chromosomes line up in the middle, right? ▷ yields to S1
52.1–55.5 **S1** [turn] Yeah, exactly. They line up single file along the metaphase plate. ▷ yields to S3
54.4–54.8 **S3** [backchannel] mm-hm
55.5–58.8 **S3** [turn] And then anaphase is when they get pulled apart, super dramatic. ▷ yields to S2
58.8–61.3 **S2** [turn] Haha, yeah, like they just zip away from each other. ▷ yields to S3
60.6–61.3 **S1** [laughter] Haha right
62.2–66.2 **S3** [turn] Okay, so do you both feel better about this stuff? Should we quiz each other real quick? ▷ yields, end

test split · tone: enthusiastic and energetic · noise: office hum, 34.3 dB SNR; keyboard clatter · voices: S1 M, S2 F, S3 M
EOT 13 yields / 0 holds · INT 3 interruptions / 7 negatives (6 backchannels, 1 laughter) · 2 failed claims (excluded)



Transcript of test_00150

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.6 **S1** [turn] Um, okay, I guess I'll start. So, uh, I pushed an update to the onboarding flow last night, mainly focusing on the animation transitions between screens. It's still a bit jumpy, but— ▷ cut off

3.9–5.1 **S3** [failed claim] Yeah, I noticed that!

9.0–10.8 **S2** [interruption] Can you clarify what you mean by jumpy? ▷ takes floor from S1

11.0–15.7 **S2** [turn] Is it like the fade between pages is too fast, or there's a lag when loading assets? ▷ yields to S1

15.6–22.8 **S1** [turn] It's more like, uh, the fade timing isn't consistent? Sometimes it just sort of snaps right over if you tap too quickly. ▷ yields to S3

23.4–28.5 **S3** [turn] Oh, right, I ran into that doing my walkthrough. Like if you double-tap next, it skips the animation. ▷ yields to S1

26.3–26.7 **S2** [backchannel] Yeah.

28.5–36.2 **S1** [turn] —that would honestly help a lot. The implementation doesn't always match the prototype, and, uh, having a GIF loop would make it clearer. ▷ yields to S3

28.6–32.8 **S3** [failed claim] Do you want me to tweak the animation speed in Figma so the devs have a clear reference?

31.7–32.1 **S2** [backchannel] mm-hm

36.2–38.6 **S3** [turn] Awesome, I'll get that together after this call. ▷ yields to S2

37.8–38.4 **S1** [backchannel] Thanks, Matt.

38.8–44.6 **S2** [turn] So, switching gears, I wanted to bring up the feedback from support about the settings menu. Users are complaining it's hidden too deep. Any ideas? ▷ cut off

43.6–48.4 **S1** [interruption] Um, maybe we could move the settings icon to the top nav? Or, like, a floating button? ▷ takes floor from S2

44.0–45.3 **S3** [backchannel] Oh, yeah, totally.

48.4–52.5 **S1** [turn] Sorry, I mean, there's not a lot of space up there, but it might make it more obvious. ▷ yields to S2

53.0–58.5 **S2** [turn] No, that's a good point. I was thinking something similar, maybe even a swipe gesture shortcut? ▷ yields to S3

58.5–60.8 **S3** [turn] Ooh, swipe from the right for settings? That's slick. ▷ cut off

59.2–59.8 **S2** [laughter] haha yeah

60.9–64.0 **S3** [interruption] Wait, but won't that interfere with back navigation on Android? ▷ takes floor from S2

64.2–66.6 **S3** [turn] I mean, maybe we could use a long-press instead? ▷ yields to S1

66.9–73.7 **S1** [turn] Yeah, or even just a persistent icon on key screens. People really just want to find their preferences faster. ▷ yields to S2

71.2–71.9 **S2** [backchannel] Exactly.

73.6–79.9 **S2** [turn] Okay, let's test a couple of those options in the prototype and send out for quick user testing by Friday. ▷ yields to S3

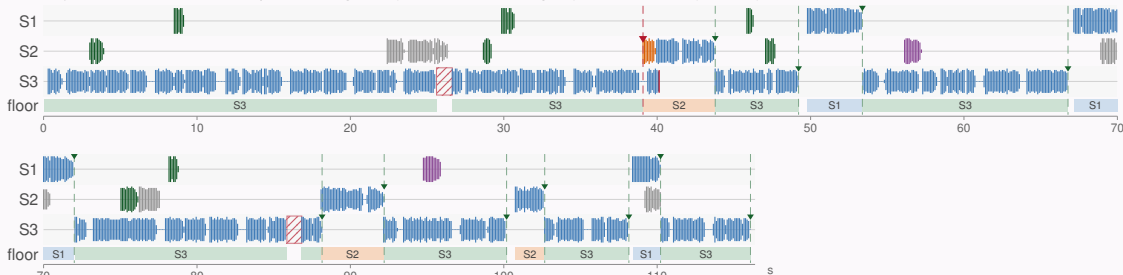
79.8–81.5 **S3** [turn] Sweet, I'm on the Figma updates. ▷ yields to S1

81.5–86.4 **S1** [turn] Uh, sounds good. Let me know if you need any assets or, um, anything from the dev side. ▷ yields to S3

84.9–85.4 **S2** [backchannel] Will do.

86.8–89.5 **S3** [turn] Great work, you two. This is really coming together! ▷ yields, end

dev split · tone: enthusiastic and energetic · noise: quiet office, 37.7 dB SNR; keyboard clatter, chairs scraping · voices: S1 F, S2 F, S3 F
EOT 12 yields / 2 holds · INT 1 interruptions / 10 negatives (8 backchannels, 2 laughter) · 4 failed claims (excluded)



Transcript of dev_00025

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–12.8 **S3** [turn] Alright, welcome! Today I'll walk you through how to approach planning a basic home renovation project, like a kitchen update. We'll focus on budgeting, scheduling, and a little on working with contractors. Ready to dive in?

3.0–3.9 **S2** [backchannel] So excited!

8.5–9.2 **S1** [backchannel] Sounds good.

13.0–25.6 **S3** [turn] Perfect! So first step, get clear about the scope. Are you just replacing cabinets and counters, or are you opening up walls? Write down a short wish list and then do a walk-through. Note every single change you're considering. ▷ holds, pauses

22.4–26.4 **S2** [failed claim] Like, even tiny stuff? New handles or, um, like faucet upgrades?

26.6–31.0 **S3** [turn] —Exactly! Everything counts, even the doorknobs, because it all affects your budget.

28.7–29.2 **S2** [backchannel] Yeah.

29.9–30.8 **S1** [backchannel] Right, makes sense.

31.1–40.1 **S3** [turn] Once you've got your list, the next thing is setting a rough budget. I always tell people, add at least fifteen percent to whatever you think it'll cost.

Unexpected stuff always pops up. > cut off

39.1–39.9 **S2** [interruption] Oh, so, wait— > takes floor from S3

40.0–43.8 **S2** [turn] If I plan for ten thousand, I should really budget like, eleven five? > yields to S3

43.8–49.2 **S3** [turn] Exactly! That cushion saves you so much stress if you find, say, old wiring behind a wall. > yields to S1

45.9–46.3 **S1** [backchannel] Mm-hm.

47.1–47.7 **S2** [backchannel] Oh, yep.

49.8–53.4 **S1** [turn] What about timelines? How do you avoid projects dragging on forever? > yields to S3

53.4–66.8 **S3** [turn] Such a good question. Start by making a super basic schedule—list each stage: demo, plumbing, electrical, everything. Pad each stage with a couple extra days. And check contractor availability way in advance. > yields to S1

56.1–57.2 **S2** [laughter] Haha, love it.

67.2–72.0 **S1** [turn] So, like—do you ever act as your own contractor, or is it always better to hire one? > yields to S3

68.9–70.4 **S2** [failed claim] Yeah, I wondered that too.

72.1–85.9 **S3** [turn] Great question. If it's a small job and you're super organized, you can coordinate trades yourself. But for anything involving permits or structural work, I really recommend hiring a pro. It's worth it for your sanity—plus, liability. > holds, pauses

75.0–76.2 **S2** [backchannel] Mm-hm, oh yeah.

76.2–77.6 **S1** [failed claim] Okay, random but—

78.2–78.8 **S1** [backchannel] So true.

86.8–88.2 **S3** [turn] —Random's good, go ahead! > yields to S2

88.1–92.2 **S2** [turn] Do you recommend online planning tools or just like, good old spreadsheets? > yields to S3

92.2–100.2 **S3** [turn] Honestly, both! I use a spreadsheet for budget stuff and Trello to track stages. But even a paper checklist works, as long as you update it. > yields to S2

94.8–95.9 **S1** [laughter] Haha, love that energy.

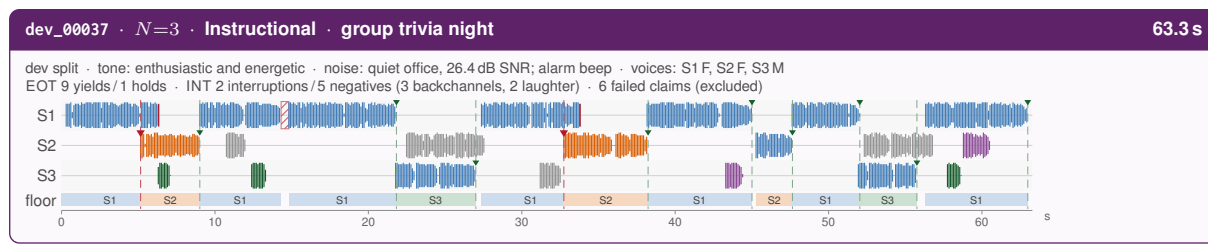
100.8–102.7 **S2** [turn] Thank you, this is super helpful. > yields to S3

102.7–108.2 **S3** [turn] My pleasure! Honestly, planning is half the battle—if you get that right, the rest is so much smoother. > yields to S1

108.4–110.2 **S1** [turn] Sounds like we're ready to actually start! > yields to S3

109.2–110.3 **S2** [failed claim] Yeah, let's do it!

110.3–116.1 **S3** [turn] Go for it! And if you run into any snags, just remember: flexibility, patience, and lots of coffee. > yields, end



Transcript of dev_00037

start–end (s) speaker [ground-truth label] text > outcome

0.3–6.3 **S1** [turn] Alright, team, this round is all about famous landmarks, so the way it works is: I'll read the question, and then you get thirty seconds to discuss before we lock in an answer. Got it? > cut off

5.1–9.0 **S2** [interruption] Wait, can we confer after you say the answer or do we have to just shout it out? > takes floor from S1 > yields to S1

6.3–7.1 **S3** [backchannel] Sounds good!

9.0–14.3 **S1** [turn] You confer first, then I collect your final answer. No shouting out until everyone's agreed. > holds, pauses

10.7–12.0 **S2** [failed claim] Yeah, that makes sense.

12.4–13.4 **S3** [backchannel] Perfect, got it.

14.8–21.8 **S1** [turn] Okay, so first question: What city is the Colosseum in? Thirty seconds on the clock. When you think you know, just talk it out. > yields to S3

21.8–23.0 **S3** [turn] That's in Rome, right?

22.5–25.6 **S2** [failed claim] Is it Rome or is it Florence? I always mix those two up—

23.1–27.0 **S3** [turn] —No, I'm pretty sure it's Rome. Florence has, like, the Duomo, not the Colosseum. > yields to S1

25.7–27.6 **S2** [failed claim] Right, right. Yeah, that rings a bell.

27.4–33.8 **S1** [turn] Exactly! Rome is the answer. See how quick that was? Okay, here's a quick tip for landmark questions: try to picture what else is nearby in each city. Helps with the tricky ones. > cut off

31.2–32.6 **S3** [failed claim] That's a good strategy.

32.7–38.2 **S2** [interruption] Wait, are we allowed to use logic outside direct memory, like, can we guess by like, what language they speak there? > takes floor from S1 > yields to S1

38.2–45.0 **S1** [turn] Absolutely! You can use any deduction you want—just no phones or outside help. It's all about your brains and teamwork. > yields to S2

43.3–44.4 **S3** [laughter] Haha, that's us.

45.3–47.6 **S2** [turn] Awesome, thanks. Okay, let's do the next one. > yields to S1

47.6–52.0 **S1** [turn] Alright! Next: Which country's flag is just a big red dot on a white background? > yields to S3

51.9–52.6 **S3** [turn] Japan!

52.3–55.4 **S2** [failed claim] Are you sure it's not Bangladesh? Don't they have a red circle?

52.7–55.8 **S3** [turn] —Their flag's got the green though, right? Japan's is just white and red. > yields to S1

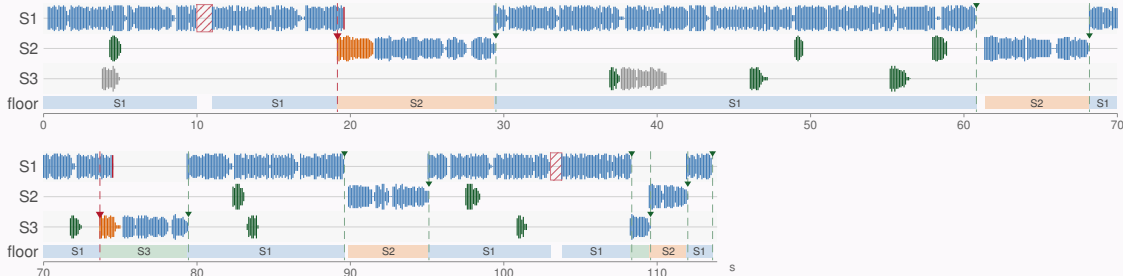
55.5–56.8 **S2** [failed claim] Yep, you're totally right.

56.3–63.0 **S1** [turn] Nailed it—Japan's the answer. You're on fire! Alright, keep using that teamwork and deduction, and you'll crush this round. > yields, end

57.8–58.7 **S3** [backchannel] Thanks, coach!

58.8–60.5 **S2** [laughter] Haha, yeah, let's keep going.

test split · tone: professional and business-like · noise: outdoor park, 34.3 dB SNR; alarm beep; distant speech · voices: S1 F, S2 F, S3 M
EOT 10 yields / 2 holds · INT 2 interruptions / 11 negatives (11 backchannels, 0 laughter) · 2 failed claims (excluded)



Transcript of test_00018

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–10.0 **S1** [turn] Good morning, everyone. So today I want to walk through the new onboarding flow for the mobile app, and just highlight some of the changes from last sprint. If you open up Figma— ▷ holds, pauses

3.8–5.0 **S3** [failed claim] Yeah, pulling it up.

4.3–5.0 **S2** [backchannel] Got it.

11.0–19.6 **S1** [turn] —you'll see we've moved the sign-up steps into a single scrollable page, instead of multiple screens. The idea is to reduce cognitive load and bounce rate by making the process feel less fragmented. ▷ cut off

19.2–21.5 **S2** [interruption] Wait, so everything's on one screen now? ▷ takes floor from S1

21.6–29.5 **S2** [turn] I'm just looking at the wireframe and trying to picture what happens if someone enters an error, like an invalid email. Does it scroll up automatically? ▷ yields to S1

29.4–42.0 **S1** [turn] Great question. We're using in-line validation, so as soon as you leave a field, it'll show a red border and an error message directly under that field. No auto-scroll, just contextual feedback right where the issue is.

36.9–37.6 **S3** [backchannel] Right.

37.7–40.6 **S3** [failed claim] And does it still support autofill for password managers?

42.1–51.1 **S1** [turn] —yes, definitely. We kept the form fields native so autofill and password management plugins still work seamlessly, especially on iOS and Android.

46.1–47.2 **S3** [backchannel] Makes sense.

49.0–49.5 **S2** [backchannel] Good.

51.2–60.8 **S1** [turn] One more thing: we've replaced the avatar upload step with a default avatar, and users can update it later from their profile. This was a bottleneck in our old flow. ▷ yields to S2

55.2–56.5 **S3** [backchannel] Yeah, much smoother.

58.0–59.0 **S2** [backchannel] I like that.

61.4–68.2 **S2** [turn] Just to clarify, is the terms and conditions checkbox mandatory on this screen? Or do we leave that to the app store permissions? ▷ yields to S1

68.2–74.5 **S1** [turn] It's mandatory—users can't create an account unless they check it. Legal reviewed the copy for us, so it's all compliant. ▷ cut off

71.7–72.5 **S3** [backchannel] Nice.

73.7–75.1 **S3** [interruption] Quick one, Jess— ▷ takes floor from S1

75.2–79.4 **S3** [turn] —in the prototype, it looks like the skip for now button is gone. Is that intentional? ▷ yields to S1

79.4–89.6 **S1** [turn] Yes, we removed the skip for now on recommendations. Analytics showed hardly anyone used it, but it complicated tracking. They can still leave sections blank and return later. ▷ yields to S2

82.3–83.1 **S2** [backchannel] Uh-huh.

83.3–84.0 **S3** [backchannel] Alright.

89.9–95.1 **S2** [turn] And you're handling localization... how, exactly? Are texts hardcoded or pulled from strings? ▷ yields to S1

95.1–103.0 **S1** [turn] It's all externalized; everything runs through the translation table. Spanish, French, and German are ready for the test build next week. ▷ holds, pauses

97.5–98.5 **S2** [backchannel] Perfect.

100.9–101.5 **S3** [backchannel] Great.

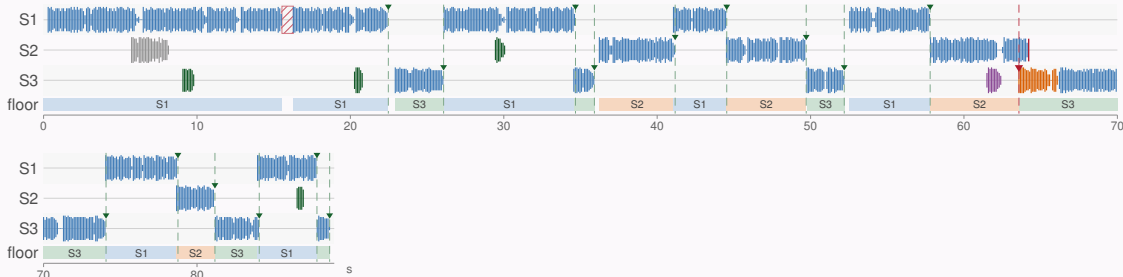
103.8–108.3 **S1** [turn] Awesome! Any last questions before we hand off to dev for the first round of QA? ▷ yields to S3

108.3–109.6 **S3** [turn] No, all clear here. ▷ yields to S2

109.5–112.0 **S2** [turn] Yeah, looks good. Thanks for running through it, Jess. ▷ yields to S1

111.9–113.6 **S1** [turn] Of course, thanks everyone! ▷ yields, end

test split · tone: enthusiastic and energetic · noise: outdoor park, 26.9 dB SNR; notification chime · voices: S1 F, S2 F, S3 F
EOT 15 yields / 1 holds · INT 1 interruptions / 5 negatives (4 backchannels, 1 laughter) · 1 failed claims (excluded)



Transcript of test_00046

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–15.5 **S1** [turn] Alright, so since we're swapping recommendations, I have to gush about "The Night Circus" by Erin Morgenstern. It's this magical realism thing where the circus just appears with no warning, and every tent is its own wonder—like, seriously, you'll want to live inside this book. ▷ holds, pauses

5.7–8.2 **S2** [failed claim] But is it slow? Sometimes those books drift for me—

9.1–9.9 **S3** [backchannel] Oh wow

16.3–22.5 **S1** [turn] —I get that! It's not exactly fast-paced, but it's super atmospheric. Like, you just wanna savor the details. > yields to S3

20.3–20.9 **S3** [backchannel] mm-hm

22.9–26.1 **S3** [turn] Does it have romance? I like a little bit, but not if it takes over. > yields to S1

26.1–34.7 **S1** [turn] It totally does, but it's subtle and woven in with all the magic stuff. It's more longing and mystery than, like, heavy-handed love triangles or anything. > yields to S3

29.5–30.1 **S2** [backchannel] Right—

34.6–35.9 **S3** [turn] That sounds perfect. > yields to S2

36.2–41.2 **S2** [turn] Okay, I'll bite. But only if you read "Arrival" the movie, I mean the short story it's based on. Ted Chiang? > yields to S1

41.1–44.5 **S1** [turn] Oh, I've seen the movie! Didn't realize it was a short story too. > yields to S2

44.5–49.7 **S2** [turn] Yeah, "Story of Your Life." The movie's good, but the story hits different. The language thing? It goes way deeper. > yields to S3

49.8–52.2 **S3** [turn] I loved "Arrival". The ending wrecked me though. > yields to S1

52.6–57.8 **S1** [turn] Same! I was ugly-crying. So, is the short story also emotional, or more sci-fi heavy? > yields to S2

57.8–64.2 **S2** [turn] It's both. Chiang's writing is super precise, but there's all this under-the-surface feeling. Like, he'll get you nerding out about linguistics and then—bam!—parenthood heartbreak. > cut off

61.5–62.4 **S3** [laughter] haha right

63.6–66.1 **S3** [interruption] Could I actually jump in with a movie pick, or— > takes floor from S2

66.2–74.1 **S3** [turn] So, "Hunt for the Wilderpeople" is this New Zealand comedy that's both hilarious and weirdly touching. It's great when you want something uplifting but, like, not cheesy. > yields to S1

74.0–78.8 **S1** [turn] Oh, I love Taika Waititi! That movie's on my list but I haven't hit play yet. > yields to S2

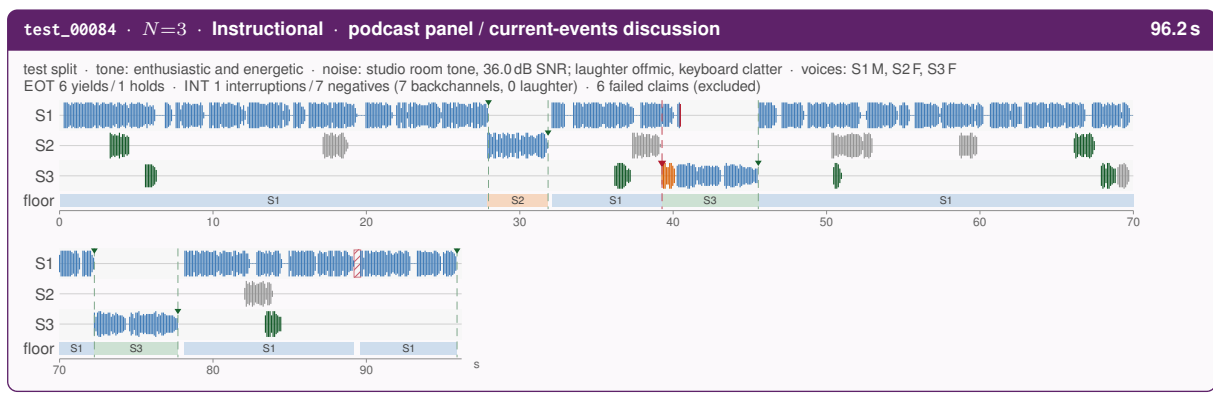
78.7–81.2 **S2** [turn] Yeah, same! Everyone keeps telling me I'd love it. > yields to S3

81.2–84.1 **S3** [turn] Trust me, it's so worth it. Ricky Baker is just—iconic. > yields to S1

84.0–87.8 **S1** [turn] Okay, looks like we've all got homework! I'm writing these down right now. > yields to S3

86.5–87.0 **S2** [backchannel] mm-hm

87.8–88.6 **S3** [turn] Me too! > yields, end



Transcript of test_00084

start-end (s) speaker [ground-truth label] text > outcome

0.3–12.1 **S1** [turn] Alright, so today I wanted to break down what's actually happening with that new carbon tax proposal everyone's talking about. There's, uh, a lot of confusion, I think, but it's not actually as wild as it sounds.

3.3–4.6 **S2** [backchannel] Oh, interesting.

5.6–6.3 **S3** [backchannel] Go on.

12.3–24.9 **S1** [turn] So, the basic idea is the government wants to put a fee, or tax, on companies for every ton of carbon dioxide they emit. It makes polluting more expensive and, in theory, encourages them to use cleaner energy.

17.2–18.8 **S2** [failed claim] So does that mean, like—

25.0–28.0 **S1** [turn] —companies that can't cut emissions just end up paying more? > yields to S2

27.9–31.8 **S2** [turn] Yeah, exactly, so would that cost get passed along to consumers then? > yields to S1

32.1–40.5 **S1** [turn] That's the worry, right? If companies pay more, a lot of analysts think some of that will show up in prices. But the policy also includes rebates—basically, refund checks for households to offset the higher costs. > cut off

36.2–37.2 **S3** [backchannel] Mm, rebates.

37.4–39.2 **S2** [failed claim] That's actually kind of cool.

39.3–40.2 **S3** [interruption] Wait, so— > takes floor from S1

40.3–45.5 **S3** [turn] Are the rebates automatic, or do people have to apply? That's always my question with these things. > yields to S1

45.6–54.6 **S1** [turn] Good question, actually. They're supposed to be automatic—like, you get it with your tax return, so you don't have to fill anything out. But, you know, implementation is everything.

50.4–53.0 **S2** [failed claim] Right, because people miss out otherwise. Uh-huh.

50.5–51.0 **S3** [backchannel] Yeah.

54.8–62.8 **S1** [turn] Exactly. And the other thing people miss is, the tax goes up a little every year, to keep up the pressure. It's not a one-time bump – it ramps.

58.7–59.9 **S2** [failed claim] Okay, but—

62.9–71.4 **S1** [turn] —yeah, if you're asking what happens to small businesses, that's a fair concern. There are carve-outs for certain sectors, like farming, so they don't get hit as hard.

66.2–67.5 **S2** [backchannel] That's reassuring.

67.9–68.9 **S3** [backchannel] Makes sense.

69.0–69.8 **S3** [failed claim] So, uh—

71.5–72.3 **S1** [turn] —go for it. > yields to S3

72.3–77.7 **S3** [turn] What's the actual impact supposed to be? Like, do experts think this'll actually cut emissions much? > yields to S1

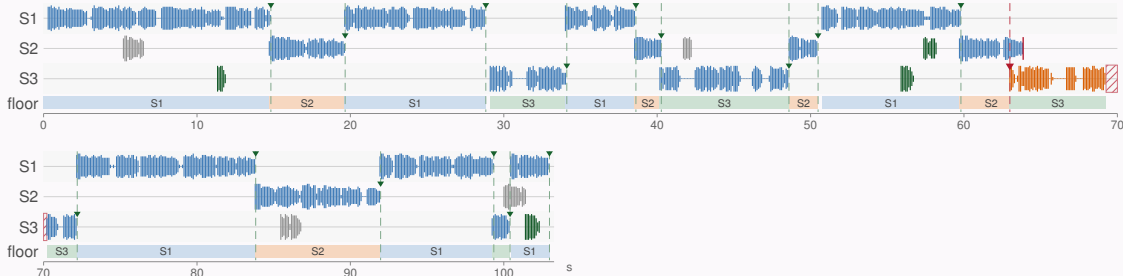
78.2–89.2 **S1** [turn] Yeah, so most studies say it definitely helps, especially if the tax keeps rising. It pushes companies to innovate. But, uh, it's not a magic bullet—needs to be part of a bigger set of policies. > holds, pauses

82.1–83.9 **S2** [failed claim] Got it. That's super helpful.

83.4–84.5 **S3** [backchannel] Yeah, thanks.

89.6–95.9 **S1** [turn] No problem, I just think it's one of those topics that sounds scary, but once you get into the details, it's kind of elegant. > yields, end

test split · tone: enthusiastic and energetic · noise: cafe, 35.3 dB SNR; alarm beep · voices: S1 M, S2 M, S3 F
EOT 15 yields / 1 holds · INT 1 interruptions / 4 negatives (4 backchannels, 0 laughter) · 4 failed claims (excluded)



Transcript of test_00122

start-end (s) speaker [ground-truth label] text > outcome

0.3–14.8 **S1** [turn] Alright, team, so we need to talk about chore rotations. The whiteboard system I'm proposing is actually super simple—it keeps everything fair. Each of us gets assigned a color and then we just rotate tasks weekly, clockwise. No one gets stuck with dishes forever. Sound good so far? > yields to S2

5.2–6.6 **S2** [failed claim] Yeah, totally makes sense.

11.3–11.9 **S3** [backchannel] Uh-huh.

14.7–19.7 **S2** [turn] So wait, if I have, like, bathroom duty this week, next week I get trash, and then laundry, or— > yields to S1

19.6–28.8 **S1** [turn] Exactly! You'll never have the same job two weeks in a row, unless someone wants to swap. I'll draw up the rotation on Sunday nights, and we can stick it on the fridge. That way we all see it. > yields to S3

29.1–34.1 **S3** [turn] Do we get to pick colors, or is this like a Sorting Hat scenario? > yields to S1

34.0–38.6 **S1** [turn] Haha, oh, you can totally pick. First come, first serve. Just don't fight over blue again. > yields to S2

38.5–40.3 **S2** [turn] Blue is elite, I stand by that. > yields to S3

40.2–44.0 **S3** [turn] Yeah, I'll take purple. I like living dangerously.

41.7–42.3 **S2** [failed claim] Wait, but—

44.1–48.6 **S3** [turn] —No, you can't have purple. Sorry, rules are rules. > yields to S2

48.6–50.5 **S2** [turn] Haha, alright, fine, I'll survive. > yields to S1

50.8–59.8 **S1** [turn] Okay, back on track—so for supplies, the person on trash duty is in charge of restocking garbage bags and paper towels. There's a little list on the side of the fridge for that. > yields to S2

55.9–56.7 **S3** [backchannel] Makes sense.

57.4–58.2 **S2** [backchannel] Oh, smart.

59.7–63.9 **S2** [turn] But like, what if, um, someone totally forgets their turn? I'm not naming names, but— > cut off

63.0–69.3 **S3** [interruption] Yeah, like when you went to Coachella last month and the kitchen became a hazardous zone— > takes floor from S2 > holds, pauses

70.2–72.2 **S3** [turn] —seriously, what happens then? > yields to S1

72.2–83.8 **S1** [turn] Honestly, if someone forgets, we just let them know. No drama, just a text—or, like, a big passive-aggressive arrow on the whiteboard. But if it keeps happening, we swap who gets to pick the next pizza place, for real consequences. > yields to S2

83.8–86.7 **S2** [turn] That's actually motivating. I need my Friday pizza.

85.5–86.8 **S3** [failed claim] The threat is real.

86.8–92.0 **S2** [turn] So, recapping—whiteboard, color-coded, rotate, pizza places at stake. Did I miss anything? > yields to S1

91.9–99.3 **S1** [turn] You nailed it, man. I'll finish setting it up tonight—you're both getting a group text with your colors. If you want stickers or something, let me know by nine. > yields to S3

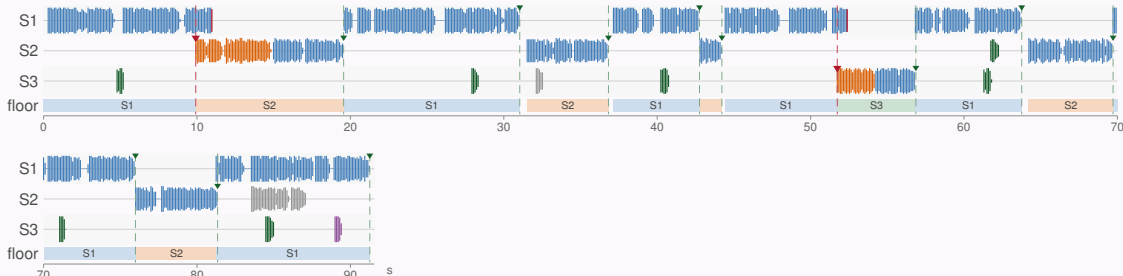
99.2–100.4 **S3** [turn] I demand glitter. > yields to S1

100.0–101.5 **S2** [failed claim] Oh wow, raising the stakes.

100.4–103.0 **S1** [turn] You got it, glitter stickers for the bravest roomie. > yields, end

101.4–102.4 **S3** [backchannel] Living the dream.

test split · tone: professional and business-like · noise: dorm room, 31.0 dB SNR; school bell · voices: S1 M, S2 M, S3 M
EOT 11 yields / 0 holds · INT 2 interruptions / 8 negatives (7 backchannels, 1 laughter) · 2 failed claims (excluded)



Transcript of test_00132

start-end (s) speaker [ground-truth label] text > outcome

0.3–11.0 **S1** [turn] Alright, let's start with Mendelian genetics since it's definitely coming up on the exam. The basic idea is that alleles segregate during gamete formation, so each gamete only gets one allele for each gene. > cut off

4.8–5.3 **S3** [backchannel] uh-huh

9.9–14.8 **S2** [interruption] Sorry, just—does that mean like, both parents pass down just one version of the gene each? > takes floor from S1

15.0–19.6 **S2** [turn] So if you have a big A and a little a, you'd only get one or the other from that parent? > yields to S1

19.6–31.0 **S1** [turn] Exactly. During meiosis, the alleles separate so the gametes get either the big A or little a, not both. When fertilization occurs, the offspring get one allele from each parent, restoring the pair. > yields to S2

27.9–28.4 **S3** [backchannel] right

31.5–33.2 **S2** [turn] Mm, okay I think I get that.

32.1–32.6 **S3** [failed claim] Wait—

33.3–36.8 **S2** [turn] —so it's like a coin flip for which one you inherit from each parent, right? ▷ yields to S1

37.2–42.8 **S1** [turn] Yeah, that's actually a good analogy. It's random. If you want, we could draw a Punnett square to visualize it— ▷ yields to S2

40.2–40.8 **S3** [backchannel] Oh, yeah.

42.8–44.2 **S2** [turn] That'd help, actually. ▷ yields to S1

44.4–52.4 **S1** [turn] Alright, so imagine a parent with genotype Aa and another with Aa. The Punnett square will have A and a on both axes, so the offspring could be AA, Aa, or aa, with the appropriate probabilities. ▷ cut off

51.7–54.2 **S3** [interruption] So, um, there's a one in four chance for getting aa? ▷ takes floor from S1

54.3–56.9 **S3** [turn] And then two out of four for Aa, and one for AA? ▷ yields to S1

56.8–63.8 **S1** [turn] That's right. It's a 1:2:1 ratio for the genotypes. The phenotypic ratio will depend on which allele is dominant. ▷ yields to S2

61.3–61.8 **S3** [backchannel] okay

61.7–62.3 **S2** [backchannel] mm-hm

64.2–69.7 **S2** [turn] Just to check, if the dominant allele is A, then both AA and Aa would show the dominant trait, right? ▷ yields to S1

69.7–76.0 **S1** [turn] Correct. Only aa would have the recessive phenotype. So it's a three to one ratio for phenotypes if A is dominant. ▷ yields to S2

71.1–71.5 **S3** [backchannel] yeah

76.0–81.3 **S2** [turn] All right, that makes sense. Thanks, I always get the genotypic and phenotypic ratios confused. ▷ yields to S1

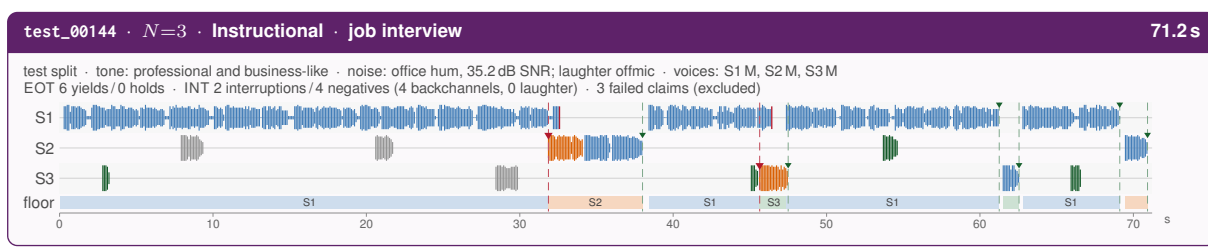
81.3–87.6 **S1** [turn] No problem, it's a common mistake. Just remember genotypes are the actual alleles while phenotypes are what you see.

83.6–87.1 **S2** [failed claim] Can we, uh, maybe go over linked genes too? Unless we need to—

84.5–85.0 **S3** [backchannel] Got it.

87.7–91.3 **S1** [turn] —No, that's great. We should definitely cover those since they're tricky. ▷ yields, end

89.0–89.5 **S3** [laughter] haha



Transcript of test_00144

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–10.7 **S1** [turn] Alright, let's get started. Today, I'm going to walk you through the structure of a standard job interview and offer some tips on how to prepare. We'll focus on behavioral questions and the STAR method. Sound good?

2.8–3.2 **S3** [backchannel] Mm-hm.

7.9–9.4 **S2** [failed claim] Yeah, that sounds good.

10.8–23.2 **S1** [turn] So interviews usually kick off with some rapport-building, maybe a question about your background. Then they'll shift to those behavioral questions—"Tell me about a time you..." situations. That's where STAR comes in: Situation, Task, Action, Result.

20.6–21.8 **S2** [failed claim] Could you, um—

23.3–32.6 **S1** [turn] —explain what that means? Right, the STAR method helps you structure your answer. First, describe the Situation, set the context. Then the Task—what you needed to achieve. After that, the Action you specifically took, and finally, the Result, what happened as a consequence. ▷ cut off

28.4–30.0 **S3** [failed claim] I always mess up the result part.

31.9–34.1 **S2** [interruption] So, like, if they ask about conflict— ▷ takes floor from S1

34.3–38.0 **S2** [turn] —would you answer using STAR too? Or is that more for, uh, achievements? ▷ yields to S1

38.4–46.4 **S1** [turn] That's a good question. STAR is for both positive and challenging situations. So with conflict, you'd outline the context, what the disagreement was, what you personally did to resolve it, and what the outcome was. ▷ cut off

45.1–45.6 **S3** [backchannel] Right.

45.6–47.5 **S3** [interruption] Can you give an example? Like, a quick one? ▷ takes floor from S1 ▷ yields to S1

47.4–61.3 **S1** [turn] Absolutely. So, let's say you managed a project and there was a deadline issue. The Situation: tight deadline. Task: deliver on time. Action: you reorganized the team's workflow, prioritized tasks. Result: submitted ahead of schedule and got positive feedback. ▷ yields to S3

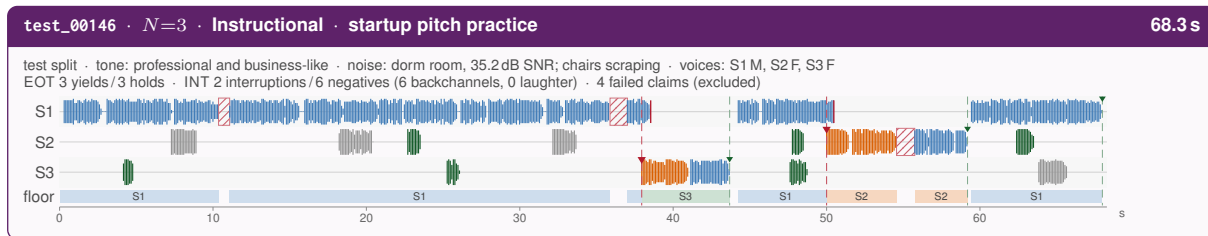
53.7–54.7 **S2** [backchannel] Oh, got it.

61.5–62.5 **S3** [turn] Yeah, that makes sense. ▷ yields to S1

62.8–69.1 **S1** [turn] Good. Also, remember to keep your answers concise, but specific. If you want, we can practice some questions right now. ▷ yields to S2

65.9–66.6 **S3** [backchannel] Yeah, sure.

69.5–70.9 **S2** [turn] Um, that'd be helpful. ▷ yields, end



Transcript of test_00146

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–10.4 **S1** [turn] Uh, alright, so, welcome everybody. Um, today I wanted to walk you through how to structure a strong startup pitch, uh, especially for a first meeting with investors. ▷ holds, pauses

4.2–4.9 **S3** [backchannel] Go on.

7.3–9.0 **S2** [failed claim] Great, I've been dying to figure this out.

11.1–20.8 **S1** [turn] Yeah, okay, so the simplest way to start is with a really clear problem statement. The more specific you are about the pain point, the easier it is for people to buy in.

18.2–20.5 **S2** [failed claim] So like, not just 'healthcare is broken,' but—

20.9–27.8 **S1** [turn] —uh, right, exactly, like ‘patients with chronic conditions wait weeks for test results,’ or something concrete like that.

22.7–23.6 **S2** [backchannel] Okay, got it.

25.3–26.1 **S3** [backchannel] Makes sense.

28.0–35.9 **S1** [turn] After you do that, you want to follow up right away with your solution. Don’t make them wait. Just one or two sentences, but show why it’s unique. ▷ holds, pauses

32.2–33.7 **S2** [failed claim] Right, like elevator pitch style?

37.0–38.5 **S1** [turn] —yeah, exactly, that’s the idea. ▷ cut off

37.9–41.0 **S3** [interruption] But what if it’s not totally unique? There’s always competition. ▷ takes floor from S1

41.2–43.7 **S3** [turn] Like, should you just say you’re better or, I don’t know, faster? ▷ yields to S1

44.2–50.5 **S1** [turn] Uh, that’s a good question. I, um, recommend highlighting what really differentiates you, even if it’s a small edge. But don’t trash-talk competitors—just be factual. ▷ cut off

47.6–48.7 **S3** [backchannel] Yeah, makes sense.

47.8–48.6 **S2** [backchannel] Oh totally.

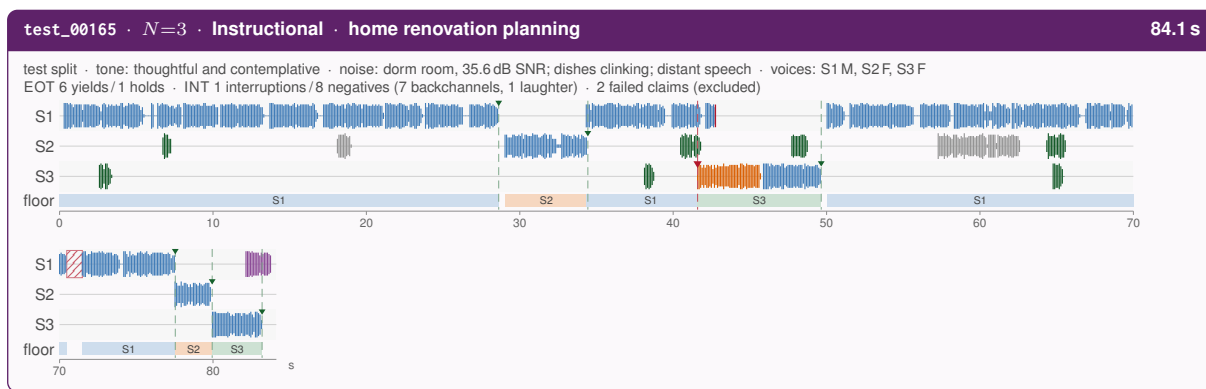
50.0–54.6 **S2** [interruption] So, when do you bring up numbers? Like, market size and projections and all that jazz? ▷ takes floor from S1 ▷ holds, pauses

55.7–59.2 **S2** [turn] I always get stuck trying to cram them in too early, or not at all. ▷ yields to S1

59.4–68.0 **S1** [turn] Yeah, it’s tempting to front-load it, but I suggest laying out the problem, then your solution, and once they’re with you, bring in the market numbers and business model. ▷ yields, end

62.4–63.6 **S2** [backchannel] Super helpful, thanks.

63.8–65.7 **S3** [failed claim] That actually clears it up a lot.



Transcript of test_00165

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–10.4 **S1** [turn] Alright, uh, so when you start planning a home renovation, the first thing you want to do is actually set your priorities. Like, what matters most to you—say, updating the kitchen versus adding more space.

2.6–3.4 **S3** [backchannel] Makes sense.

6.7–7.3 **S2** [backchannel] Right.

10.5–21.1 **S1** [turn] Once you’ve made that list, uh, you want to think about your budget. And I mean a realistic number, not just what you wish you could spend. Include a bit extra for, like, unexpected stuff that almost always comes up.

18.1–19.0 **S2** [failed claim] So, um—

21.3–28.6 **S1** [turn] —for example, if your estimate is fifty thousand, I’d set aside at least ten percent more for overages. It really saves you stress down the road. ▷ yields to S2

29.0–34.4 **S2** [turn] How do you even make a good estimate if you have no clue about costs? Like, is there a starting point? ▷ yields to S1

34.3–42.8 **S1** [turn] Yeah, uh, usually you’d look up average costs in your region or talk to contractors for rough quotes. There are websites with calculators too, but nothing beats getting at least two or three actual professionals to look at what you want to do. ▷ cut off

38.1–38.7 **S3** [backchannel] Got it.

40.5–41.8 **S2** [backchannel] Yeah, that’s helpful.

41.6–45.7 **S3** [interruption] But, like, how do you know if a contractor’s actually reliable? I mean, what if they lowball? ▷ takes floor from S1

45.9–49.6 **S3** [turn] I’ve heard horror stories about people taking deposits and just disappearing. ▷ yields to S1

47.7–48.8 **S2** [backchannel] Yeah, totally.

50.0–62.0 **S1** [turn] That’s a fair concern. Uh, so you want to check their licensing, read reviews, and absolutely call their references. If anything feels off, just walk away. And never pay more than, say, ten percent upfront. Like, stagger your payments.

57.3–62.6 **S2** [failed claim] And what about, um, permits? Who’s actually in charge of that? Is it, like, always the contractor, or—

62.1–70.5 **S1** [turn] —good question. Usually, the contractor applies, but you want to clarify that before you start. Because if you skip permits, you could get fined or—worse—have to undo the work. ▷ holds, pauses

64.4–65.6 **S2** [backchannel] That’s a nightmare.

64.8–65.5 **S3** [backchannel] Oh, wow.

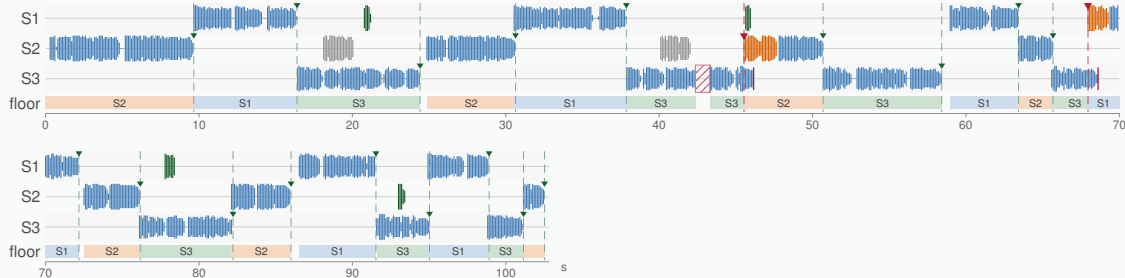
71.5–77.5 **S1** [turn] Exactly. So, just make sure everything’s above board—ask questions, get it all in writing, and, uh, don’t rush any decisions. ▷ yields to S2

77.5–79.9 **S2** [turn] Thanks, that’s honestly really reassuring. ▷ yields to S3

80.0–83.2 **S3** [turn] Yeah, I feel like I might actually survive this process now. Haha. ▷ yields, end

82.2–83.8 **S1** [laughter] Haha, you’ll both be fine.

dev split · tone: professional and business-like · noise: living room tv distant, 21.5 dB SNR; school bell, door knock · voices: S1 F, S2 F, S3 M
EOT 18 yields / 1 holds · INT 2 interruptions / 4 negatives (4 backchannels, 0 laughter) · 2 failed claims (excluded)



Transcript of dev_00044

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.7 **S2** [turn] Alright, team, we need to finalize the restaurant for the client dinner next week. I know Italian's been floated—any strong objections before I call and reserve? ▷ yields to S1

9.7–16.4 **S1** [turn] Well, I'm not sure Italian's really the safest bet—what about dietary restrictions? Michael, didn't you say someone was vegan? ▷ yields to S3

16.4–19.2 **S3** [turn] Yeah, uh, Janet from their side—she's vegan, but—

18.1–20.1 **S2** [failed claim] But most places can do a pasta, right?

19.3–24.4 **S3** [turn] ,but I don't think she'll want a salad and plain noodles again, she actually mentioned that at the last event. ▷ yields to S2

20.8–21.2 **S1** [backchannel] Right

24.9–30.6 **S2** [turn] No, that's a good point. So are we pivoting to something more, um, flexible? Mediterranean maybe? ▷ yields to S1

30.6–37.9 **S1** [turn] I mean, Mediterranean's usually good for vegetarians, but full vegan, it still depends on the spot. Have you looked into any menus, Michael? ▷ yields to S3

37.9–42.4 **S3** [turn] Hold on—I did. I pulled up two places; Petra has lots of vegan mezzes, but— ▷ holds, pauses

40.1–42.0 **S2** [failed claim] Isn't Petra that noisy place on Ninth?

43.3–46.2 **S3** [turn] ,yes, but if we reserve the back room, it's quiet enough, I checked reviews just now. ▷ cut off

45.5–47.7 **S2** [interruption] How's their wine list? Clients are big on that— ▷ takes floor from S3

45.6–46.0 **S1** [backchannel] mm-hm

47.8–50.7 **S2** [turn] —I just don't want to get caught with two reds and nothing else. ▷ yields to S3

50.7–58.4 **S3** [turn] It's surprisingly broad, actually. They have Lebanese whites and a Greek rosé that got good reviews. Want me to send the wine list over right now? ▷ yields to S1

59.0–63.4 **S1** [turn] Yes, please. Also—did anyone check allergies? Last time we had a real mess with nuts. ▷ yields to S2

63.4–65.7 **S2** [turn] Good catch, Kelly's allergic to pine nuts, right? ▷ yields to S3

65.6–68.6 **S3** [turn] Exactly. Petra uses pine nuts in some dishes— ▷ cut off

68.0–69.3 **S1** [interruption] So that's a problem— ▷ takes floor from S3

69.4–72.2 **S1** [turn] —unless they can make everything nut-free for our table. ▷ yields to S2

72.5–76.2 **S2** [turn] Should I just call them and ask? Or do we need to look at a different option? ▷ yields to S3

76.1–82.2 **S3** [turn] Wait, let's not bail too fast. I know the manager there—I can email and ask directly about vegan and nut-free. ▷ yields to S2

77.8–78.5 **S1** [backchannel] oh great

82.1–86.0 **S2** [turn] Perfect, can you copy us in? If they confirm both, I'm good reserving. ▷ yields to S1

86.5–91.5 **S1** [turn] Mm, but what about location—some clients are coming from uptown, isn't Petra kind of far south? ▷ yields to S3

91.6–95.0 **S3** [turn] I thought about that—it's only a ten-minute cab from their hotel. ▷ yields to S1

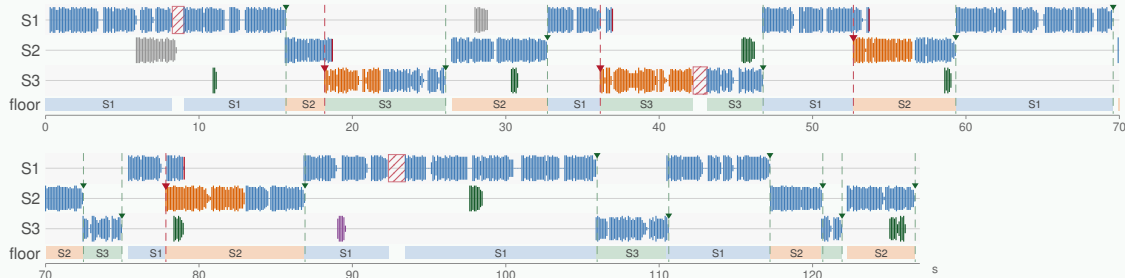
93.0–93.4 **S2** [backchannel] Right

94.9–98.9 **S1** [turn] Alright, so pending your email, Petra it is? As long as we get those answers. ▷ yields to S3

98.8–101.2 **S3** [turn] Yep, I'll send it out right after this call. ▷ yields to S2

101.2–102.5 **S2** [turn] Sounds good, thanks everyone. ▷ yields, end

dev split · tone: thoughtful and contemplative · noise: office hvac, 21.2 dB SNR; notification chime; distant speech · voices: S1 M, S2 M, S3 M
EOT 15 yields / 3 holds · INT 4 interruptions / 8 negatives (7 backchannels, 1 laughter) · 2 failed claims (excluded)



Transcript of dev_00048

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.3 **S1** [turn] Alright, so I think if we start with the probability chapter, that's the one everyone had trouble with, especially the conditional stuff. ▷ holds, pauses

5.9–8.5 **S2** [failed claim] Yeah, especially like that weird coin problem with the—

9.0–15.7 **S1** [turn] —the two coins and you pick one without knowing which? Exactly. So, okay, let's set up the variables first. ▷ yields to S2

10.9–11.2 **S3** [backchannel] mm

15.6–18.7 **S2** [turn] Alright so you've got coin A, which is fair, and coin B, which has heads on both sides, right? ▷ cut off

18.2–21.9 **S3** [interruption] Wait, but is it heads on both for B, or tails on both? ▷ takes floor from S2

22.0–26.1 **S3** [turn] 'Cause the practice one from class had two heads, so I think, yeah. ▷ yields to S2

26.5–29.3 **S2** [turn] Oh, right, yeah, it was heads. So, if you pick one and get heads—

28.0–28.9 **S1** [failed claim] But hang on—

29.4–32.7 **S2** [turn] —the question's what's the probability it was the double-heads coin, right? ▷ yields to S1

30.4–30.8 **S3** [backchannel] yeah

32.8–37.0 **S1** [turn] Right, so you basically want $P(\text{it's coin B} \mid \text{you saw heads})$. That's Bayes' theorem. ▷ cut off

36.2–42.2 **S3** [interruption] Wait, uh, do you set up the numerator as just the probability you pick B times the chance you get heads? ▷ takes floor from S1 ▷ holds, pauses

43.1–46.8 **S3** [turn] So like one-half times one, 'cause it's always heads for B? ▷ yields to S1

45.4–46.3 **S2** [backchannel] yeah, exactly

46.8–53.7 **S1** [turn] Yeah, so the numerator is one-half, but don't forget the denominator is the total probability of getting heads, so you need both coins. For coin A, it's one-half chance of heads. ▷ cut off

52.7–56.6 **S2** [interruption] So is that like, one-half times one for B plus one-half times one-half for A? ▷ takes floor from S1

56.7–59.3 **S2** [turn] So... one-half plus one-fourth, which is three-fourths? ▷ yields to S1

58.8–59.1 **S3** [backchannel] uh-huh

59.4–69.6 **S1** [turn] So then the answer would be one-half divided by three-fourths, which comes out to two-thirds. So there's a two-thirds chance you picked the double-heads coin if you saw heads. ▷ yields to S2

69.9–72.5 **S2** [turn] That's always so counterintuitive to me, honestly. ▷ yields to S3

72.4–75.0 **S3** [turn] I mean yeah it feels wrong every time. ▷ yields to S1

75.4–79.0 **S1** [turn] The trick is just to build the tree, like—okay, actually, let me draw it out. ▷ cut off

77.8–83.0 **S2** [interruption] Can we do like, the disease test example right after? That's the other one I keep getting tripped up on. ▷ takes floor from S1

78.4–79.1 **S3** [backchannel] oh right

83.1–86.9 **S2** [turn] Like with the false positives? 'Cause I always mix up when to use the base rate. ▷ yields to S1

86.8–92.3 **S1** [turn] Yeah, totally, let's do that next. And it's literally the same logic, just different numbers. ▷ holds, pauses

89.0–89.6 **S3** [laughter] haha

93.5–106.0 **S1** [turn] Okay, so for the tree: first branch splits by which coin you drew, then each branch splits for heads or tails. That's what helps me visualize the probabilities, otherwise the fractions just get lost. ▷ yields to S3

97.7–98.5 **S2** [backchannel] oh definitely

105.9–110.6 **S3** [turn] Yeah, I mean, I just write random boxes with arrows until something makes sense. ▷ yields to S1

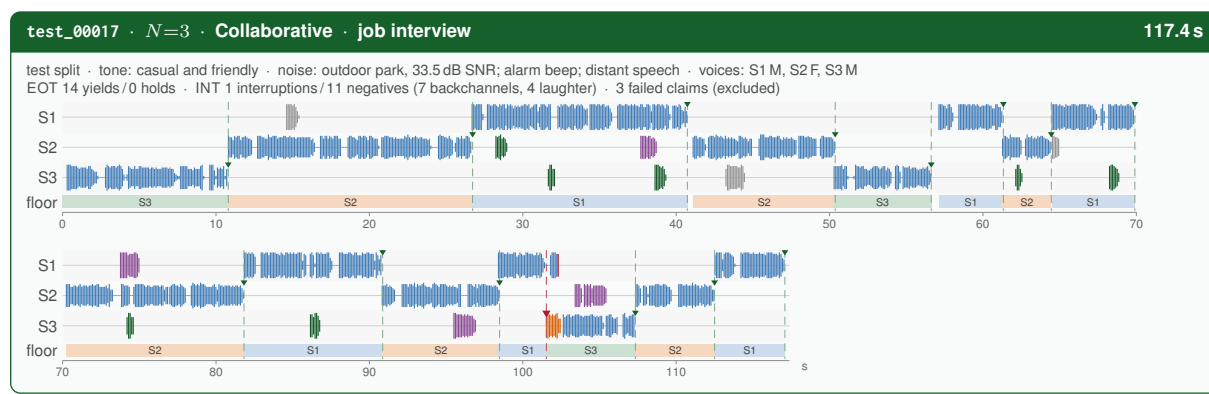
110.5–117.2 **S1** [turn] Honestly whatever gets you to the answer, right? I guess so—okay, who wants to set up the disease test one? ▷ yields to S2

117.3–120.7 **S2** [turn] I'll do it, but you have to double-check my math so I don't make a fool of myself. ▷ yields to S3

120.6–121.9 **S3** [turn] Spoiler: we will. ▷ yields to S2

122.3–126.7 **S2** [turn] Thanks. So, say the disease is rare, one percent, and test is ninety-nine percent accurate... ▷ yields, end

125.0–126.1 **S3** [backchannel] classic setup



Transcript of test_00017

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–10.8 **S3** [turn] Alright, thanks for coming in, both of you. So, to get us started, maybe you can each share a little about how you approach teamwork, especially when things get, uh, a little hectic? ▷ yields to S2

10.8–20.6 **S2** [turn] Ooh, I love this question! Honestly, I think the best things happen under a tiny bit of chaos. Like, I kind of thrive on it? There was this one project last year—

14.6–15.4 **S1** [failed claim] Can I just—

20.8–26.7 **S2** [turn] —we were launching a campaign and everything was going sideways. And that's when, like, you just have to start— ▷ yields to S1

26.7–34.3 **S1** [turn] Yeah, sorry, I was just going to say that resonates with me because I actually think clear roles help but sometimes you have to improvise.

28.3–29.0 **S2** [backchannel] Exactly!

31.7–32.2 **S3** [backchannel] right

34.4–40.7 **S1** [turn] You know, like, on my last team, we literally had a communication channel called 'all hands on deck' for those moments. ▷ yields to S2

37.7–38.7 **S2** [laughter] Haha, that's awesome.

38.6–39.3 **S3** [backchannel] I like that.

41.1–47.9 **S2** [turn] But sometimes, you know, if everyone's talking at once, nothing moves. So I always try to be the, like, person who says—

43.2–44.5 **S3** [failed claim] Let's, let's bring it back—

48.0–50.4 **S2** [turn] —“Okay, let's make a list!” Sorry, go ahead. ▷ yields to S3

50.3–56.6 **S3** [turn] No, that's great! I was wondering, do you ever find it tough to, uh, speak up when other people are louder? ▷ yields to S1

57.1–61.3 **S1** [turn] Definitely, yeah. Early on in my career, I was super hesitant to jump in. ▷ yields to S2

61.3–64.5 **S2** [turn] Same! I used to, like, second guess myself all the time. ▷ yields to S1

62.1–62.6 **S3** [backchannel] mm-hm

64.5–69.9 **S1** [turn] —I found that if you prep even a little, you can be more confident. Like, jotting down points beforehand. ▷ yields to S2

64.5–65.0 **S2** [failed claim] But—

68.3–68.9 **S3** [backchannel] Totally.

70.2–76.9 **S2** [turn] Yeah, and just not worrying so much about being perfect. Honestly, some of my best ideas have come out of mistakes.

73.8–75.1 **S1** [laughter] Haha, that's reassuring.

74.2–74.7 **S3** [backchannel] right

77.0–81.8 **S2** [turn] Do you guys, um, do anything special to encourage quieter team members to join in? ▷ yields to S1

81.8–90.9 **S1** [turn] Actually, yeah. I, uh, try to open a space by, like, asking directly—just 'Hey, what do you think?' even if it means a bit of silence while they gather their thoughts. ▷ yields to S2

86.2–86.9 **S3** [backchannel] Oh, yeah!

90.8–98.5 **S2** [turn] I do the same. Or sometimes I send a chat message if I can tell they want to say something but, you know, the group is a little too rowdy. ▷ yields to S1

95.5–97.0 **S3** [laughter] Haha, that sounds familiar.

98.4–102.3 **S1** [turn] Sometimes it's also about, uh, modeling when to step back. I'll literally say, 'I think I've talked enough, someone else jump in!' **▷ cut off**

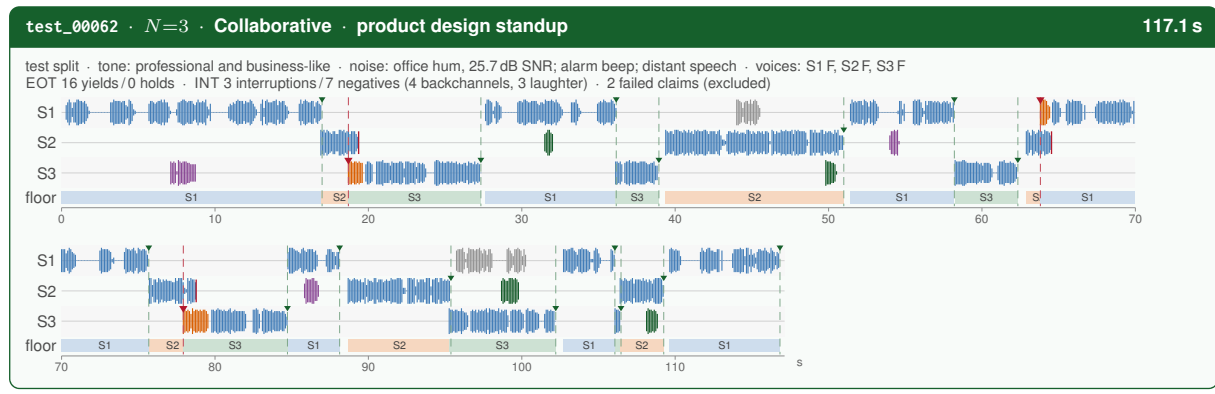
101.5–102.5 **S3** [interruption] So, real quick— **▷ takes floor from S1**

102.6–107.3 **S3** [turn] —if you had to describe your leadership style in, like, a sentence, what would you say? **▷ yields to S2**

103.4–105.5 **S2** [laughter] Haha, I wish more people would do that.

107.4–112.5 **S2** [turn] Oh! Uh, enthusiastic delegator. I make sure everyone's involved and having fun. **▷ yields to S1**

112.5–117.1 **S1** [turn] Supportive facilitator. I try to bring out the best in everyone, no matter their style. **▷ yields, end**



Transcript of test_00062

start–end (s) speaker [ground-truth label] text **▷** outcome

0.3–17.0 **S1** [turn] Okay, is everyone here? I'll just kick us off—um, main agenda item is the onboarding flow update, so let's try not to spiral into existential despair about color palettes again. **▷ yields to S2**

7.1–8.8 **S3** [laughter] Haha, I make no promises.

16.9–19.4 **S2** [turn] I think we're all good. So, about onboarding— **▷ cut off**

18.7–19.7 **S3** [interruption] Yeah, can I just— **▷ takes floor from S2**

19.8–27.3 **S3** [turn] Sorry, I wanted to ask if everyone saw the Figma I dropped last night? I added those little micro-interactions to the progress dots. **▷ yields to S1**

27.6–36.2 **S1** [turn] I did, yeah. The confetti explosion when you finish onboarding was—subtle. Extremely subtle. **▷ yields to S3**

31.5–32.1 **S2** [backchannel] mm-hm

36.1–38.9 **S3** [turn] I mean, why not celebrate the little victories, right? **▷ yields to S2**

39.4–46.2 **S2** [turn] Actually, yeah, I think the animation does make it feel a bit more welcoming. Do we feel like it's too, um, distracting or—

44.0–45.5 **S1** [failed claim] —Only if you hate joy.

46.3–51.0 **S2** [turn] —or should we tone it down for accessibility? Like motion sensitivity settings? **▷ yields to S1**

49.8–50.5 **S3** [backchannel] oh totally

51.4–58.2 **S1** [turn] We could make the confetti optional. Like, opt in if you want your eyes assaulted. **▷ yields to S3**

54.0–54.6 **S2** [laughter] Haha

58.2–62.3 **S3** [turn] I could add a toggle for reduced motion, for sure. That's actually overdue. **▷ yields to S2**

62.9–64.5 **S2** [turn] Great, thanks. So, uh, copy— **▷ cut off**

63.8–64.5 **S1** [interruption] Can I— **▷ takes floor from S2**

64.6–75.7 **S1** [turn] Sorry, I just wanted to check if we're still keeping that step-by-step tooltip intro. Engineering said it's a pain to maintain. **▷ yields to S2**

75.7–78.8 **S2** [turn] Good point. Um, do we have usage stats on it? If no one's reading those, maybe— **▷ cut off**

77.9–79.6 **S3** [interruption] Actually, I checked Mixpanel— **▷ takes floor from S2**

79.8–84.7 **S3** [turn] Barely anyone clicks past the first tooltip. It's like, 9% completion rate. **▷ yields to S1**

84.8–88.1 **S1** [turn] Staggeringly high engagement, as always. **▷ yields to S2**

85.8–86.8 **S2** [laughter] haha right

88.7–95.4 **S2** [turn] Honestly, should we just scrap the tooltips and do a single welcome banner? Would that make onboarding, uh, feel less hand-holdy? **▷ yields to S3**

95.3–96.5 **S3** [turn] I vote yes. I mean—

95.7–100.3 **S1** [failed claim] Or maybe a modal nobody reads. Very on brand.

96.7–102.2 **S3** [turn] —users basically X out anything that pops up, so maybe just... nothing flashy? Keep it simple? **▷ yields to S1**

98.7–99.8 **S2** [backchannel] yeah exactly

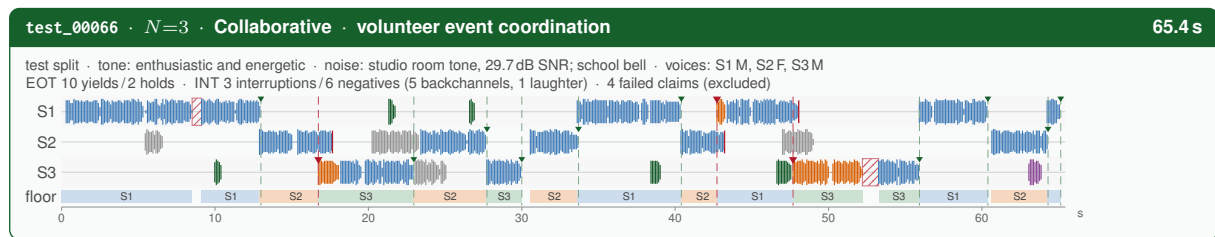
102.7–106.1 **S1** [turn] All in favor of minimalism say 'aye.' **▷ yields to S3**

106.1–106.5 **S3** [turn] Aye! **▷ yields to S2**

106.4–109.3 **S2** [turn] Aye. I do think we'll need to run this by product, though. **▷ yields to S1**

108.1–108.9 **S3** [backchannel] yeah for sure

109.7–116.8 **S1** [turn] Absolutely. Let's schedule a sync, and we can all practice our persuasive faces. **▷ yields, end**



Transcript of test_00066

start–end (s) speaker [ground-truth label] text **▷** outcome

0.3–8.5 **S1** [turn] Um, so I was thinking maybe we could, uh, meet at the community center an hour early? Just to, like, set everything up before folks arrive. **▷ holds, pauses**

5.5–6.6 **S2** [failed claim] That's a good start—

9.1–13.0 **S1** [turn] —uh, and maybe sort out the tables, I mean, so volunteers know where to go? **▷ yields to S2**

10.0–10.5 **S3** [backchannel] yeah

12.9–17.7 **S2** [turn] But let's make sure we're not just standing around. I wanna assign people roles before they walk in. Otherwise it gets chaotic. > cut off

16.7–18.1 **S3** [interruption] Um, do we have a list— > takes floor from S2

18.2–23.0 **S3** [turn] Do we have a list of who's coming? Because I can, uh, print name tags tonight if that helps. > yields to S2

20.2–23.3 **S2** [failed claim] We should. I think Josh emailed it but it's probably out of date—

21.3–21.8 **S1** [backchannel] right

23.0–25.1 **S3** [failed claim] —I can double-check the sign-ups, uh—

23.4–27.7 **S2** [turn] I'll forward you the spreadsheet, Kyle, so you get the latest names. We don't want to miss anyone. > yields to S3

26.6–27.0 **S1** [backchannel] mm-hm

27.7–30.0 **S3** [turn] Thanks! Uh, I'll get those printed first thing. > yields to S2

30.6–33.7 **S2** [turn] Also, food. Did anyone talk to that bakery about the muffins? > yields to S1

33.6–40.4 **S1** [turn] Yeah, uh, I called them but they said they'd only confirm the order if we, uh, give them a headcount by noon tomorrow. > yields to S2

38.4–39.1 **S3** [backchannel] oh wow

40.4–43.2 **S2** [turn] So we have to finalize the RSVPs tonight, basically. Otherwise, no muffins. > cut off

42.7–43.3 **S1** [interruption] If, um— > takes floor from S2

43.4–48.1 **S1** [turn] If we need help following up, I can send reminder texts to the late responses? > cut off

46.6–47.6 **S3** [backchannel] good idea

47.0–49.0 **S2** [failed claim] Yes, do that, please! Oh, and can someone—

47.7–52.2 **S3** [interruption] —uh, should we put the floor plan in the WhatsApp group? So people see it before they show up? > takes floor from S2 > holds, pauses

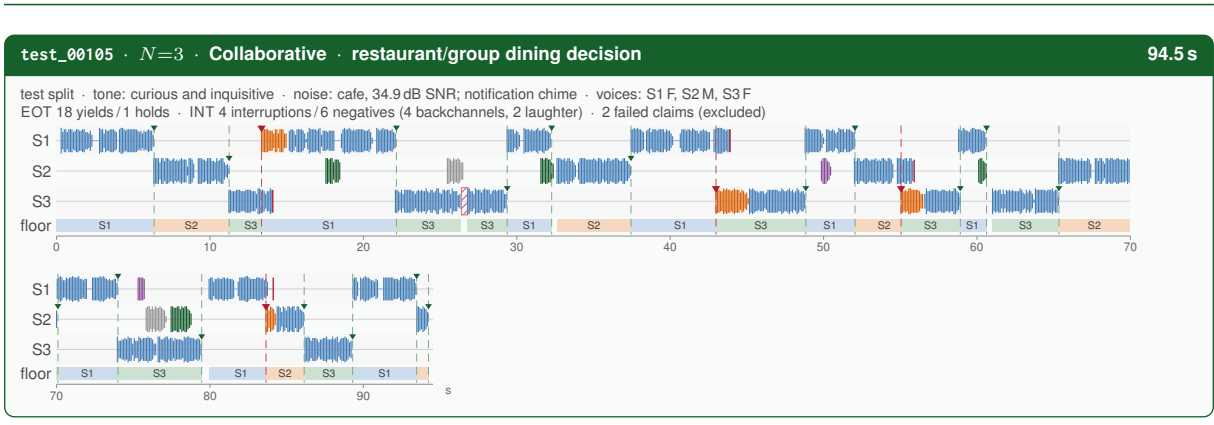
53.3–55.9 **S3** [turn] I can, um, make a quick diagram if that helps. > yields to S1

55.9–60.4 **S1** [turn] That would be awesome. Uh, just so everyone knows where to, um, set up when they arrive. > yields to S2

60.6–64.3 **S2** [turn] Perfect, let's keep the momentum. We have, like, what—two days? Let's nail it, people! > yields to S1

63.1–63.9 **S3** [laughter] haha yeah

64.2–65.1 **S1** [turn] Yeah, we got this. > yields, end



Transcript of test_00105

start–end (s) speaker [ground-truth label] text > outcome

0.3–6.4 **S1** [turn] Okay! So, where are we eating, people? I am officially starving and I might gnaw my own arm if we keep debating. > yields to S2

6.3–11.2 **S2** [turn] Uh, right, um, well I, I mean, maybe something like Thai? Unless, unless everyone's sick of noodles. > yields to S3

11.2–14.1 **S3** [turn] We just did Thai last week though, didn't we? I'm honestly kind of craving tacos. > cut off

13.4–15.0 **S1** [interruption] Oh tacos! Wait, hold up— > takes floor from S3

15.2–22.2 **S1** [turn] So, hear me out, what if we go to that fusion place? The one with the bulgogi tacos and like, everything else under the sun? > yields to S3

17.6–18.5 **S2** [backchannel] Oh yeah, uh-huh

22.1–26.4 **S3** [turn] Fusion's okay, but wasn't it super crowded last time? I had to stand for like ten minutes. > holds, pauses

25.5–26.6 **S2** [failed claim] Yeah, true, but—

26.8–29.4 **S3** [turn] —they do have great guac so, honestly, maybe worth it? > yields to S1

29.4–32.3 **S1** [turn] Haha, okay, I'd legit wait forever for that guac. > yields to S2

31.6–32.4 **S2** [backchannel] Yeah, same.

32.6–37.4 **S2** [turn] But uh, what about pricing? I, I mean, is the fusion place, like, actually kind of expensive? > yields to S1

37.5–43.9 **S1** [turn] It is, but like, not ruin-your-wallet pricey. Could we split a couple big plates maybe, instead of everyone getting their own? > cut off

43.0–45.1 **S3** [interruption] Honestly I'd rather just get my own, but— > takes floor from S1

45.2–48.8 **S3** [turn] If everyone's cool with splitting, I guess I'll survive. I'm picky but I'll deal. > yields to S1

48.8–52.0 **S1** [turn] Haha you are so dramatic, it's food, not a group project. > yields to S2

49.9–50.5 **S2** [laughter] haha

52.0–55.9 **S2** [turn] But um, can we, like, just check that it's open tonight? I think they close early sometimes. > cut off

55.1–56.5 **S3** [interruption] Oh—wait, I can look it up— > takes floor from S2

56.6–58.9 **S3** [turn] —pretty sure they're open but let me double-check. > yields to S1

58.8–60.6 **S1** [turn] Thanks boss! You're our food detective. > yields to S3

60.1–60.7 **S2** [backchannel] mm-hm

61.0–65.3 **S3** [turn] Okay, website says they're open till ten. We're good. But if there's a wait, what's our backup? > yields to S2

65.3–70.1 **S2** [turn] Maybe that Korean BBQ across the street? It's, um, kind of fun to grill your own stuff. > yields to S1

70.0–74.0 **S1** [turn] Love it! But do they have veggie options? You know, for our resident picky eater? > yields to S3

74.0–76.5 **S3** [turn] Excuse me, I'm not picky, I just know what I like.

75.3–75.8 **S1** [laughter] haha

75.8–77.2 **S2** [failed claim] Right, okay, so—

76.6–79.5 **S3** [turn] —they do have veggie, like, tofu and stuff. I've been before. > yields to S1

77.5–78.8 **S2** [backchannel] That's good, okay.

79.9–84.1 **S1** [turn] So, fusion if there's no line, Korean BBQ if it's a madhouse. And if both are packed— > cut off

83.7–84.3 **S2** [interruption] Pizza? > takes floor from S1

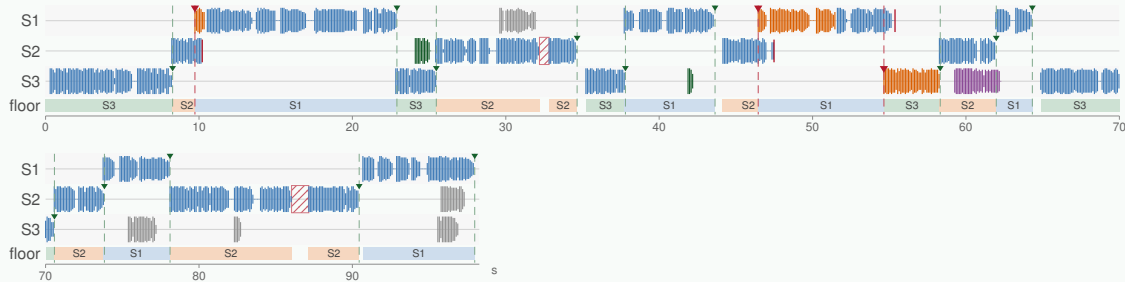
84.4–86.1 **S2** [turn] Just, just putting it out there. > yields to S3

86.2–89.3 **S3** [turn] I'll never say no to pizza, but please let's try fusion first. > yields to S1

89.3–93.5 **S1** [turn] Deal. I'm grabbing my keys. If someone doesn't move, I'm going without you. > yields to S2

93.5–94.2 **S2** [turn] Let's go! > yields, end

test split · tone: casual and friendly · noise: quiet apartment, 27.9 dB SNR; keyboard clatter, chairs scraping · voices: S1 F, S2 M, S3 F
EOT 14 yields / 2 holds · INT 3 interruptions / 3 negatives (2 backchannels, 1 laughter) · 5 failed claims (excluded)



Transcript of test_00152

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.3 **S3** [turn] Wait, so for the stats chapter, are we supposed to know how to calculate the variance formula by hand? Or do you think she'll just give us that on the test? ▷ yields to S2

8.2–10.2 **S2** [turn] I think she expects us to actually derive it, like— ▷ cut off

9.7–10.4 **S1** [interruption] But hold on— ▷ takes floor from S2

10.5–22.9 **S1** [turn] I was looking at last year's exam and she just wrote, like, 'define variance' and there was no formula included. So I feel like we really do need to memorize the steps. Not just, uh, plug and chug. ▷ yields to S3

22.8–25.5 **S3** [turn] Right, so it's like, the mean of the squared differences, right? ▷ yields to S2

24.1–25.1 **S2** [backchannel] Yeah, exactly.

25.4–32.2 **S2** [turn] You subtract the mean from each data point, square that, add 'em up, and then divide by n or n minus one, depending. ▷ holds, pauses

29.6–32.0 **S1** [failed claim] Well, for a sample it's n minus one, right?

32.8–34.6 **S2** [turn] —right, for population it's just n. ▷ yields to S3

35.2–37.8 **S3** [turn] So if she doesn't specify, should we assume it's sample? ▷ yields to S1

37.7–43.6 **S1** [turn] I think so. She usually phrases it like, 'given this sample', so we play it safe with n minus one. ▷ yields to S2

41.9–42.2 **S3** [backchannel] mm-hm

44.2–47.5 **S2** [turn] And remember the cheat sheet can only be one index card, so should we put that formula at the top? ▷ cut off

46.5–51.5 **S1** [interruption] But wait, I thought we weren't allowed to bring formula sheets at all? Or did she change her mind? ▷ takes floor from S2

51.6–55.4 **S1** [turn] Because, like, she said open note at first but then someone asked and she said just pens and calculators. ▷ cut off

54.7–58.3 **S3** [interruption] Oh no, really? I totally thought it was the card thing. Ugh, I better double— ▷ takes floor from S1 ▷ yields to S2

58.3–62.0 **S2** [turn] Actually, the email from Monday said no notes, just calculators. ▷ yields to S1

59.3–62.3 **S3** [laughter] haha oh my god I almost walked in with, like, a whole notebook.

61.9–64.3 **S1** [turn] Same, honestly. Glad you said something. ▷ yields to S3

64.9–70.6 **S3** [turn] So for probability, do you guys remember that tree diagram from lecture? The one with the colored balls? ▷ yields to S2

70.6–73.8 **S2** [turn] With the red and blue balls, and you pick two without replacement? ▷ yields to S1

73.8–78.1 **S1** [turn] Yeah, and like, if you get a red first, your chances change for the second pick. ▷ yields to S2

75.4–77.2 **S3** [failed claim] Right, conditional probability stuff.

78.1–86.0 **S2** [turn] So you'd draw out the branches for each outcome, multiply along the branches, and then sum if it's, like, 'what's the chance of getting at least one red'. ▷ holds, pauses

82.3–82.8 **S3** [failed claim] Wait—

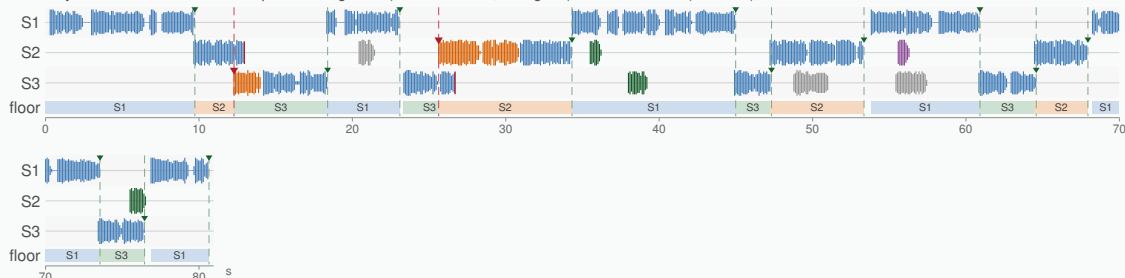
87.2–90.4 **S2** [turn] —I always get tripped up on whether to add or multiply, honestly. ▷ yields to S1

90.7–98.0 **S1** [turn] If it's 'and', you multiply; if it's 'or', you add. But just make sure the events are independent before multiplying. ▷ yields, end

95.6–96.9 **S3** [failed claim] That's super helpful, thanks!

95.8–97.4 **S2** [failed claim] Yeah, that makes it way clearer.

test split · tone: professional and business-like · noise: quiet apartment, 29.6 dB SNR; laughter offmic · voices: S1 M, S2 M, S3 M
EOT 13 yields / 0 holds · INT 2 interruptions / 4 negatives (3 backchannels, 1 laughter) · 3 failed claims (excluded)



Transcript of test_00161

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.7 **S1** [turn] Alright, thanks for joining, both of you. Let's get started on the park cleanup logistics for Saturday—first, do we have a sense of the volunteer count? ▷ yields to S2

9.7–13.0 **S2** [turn] Uh, last I checked, we had thirteen sign-ups, but it might have gone up a little since then. ▷ cut off

12.3–14.1 **S3** [interruption] Actually, I, uh, just got a text— ▷ takes floor from S2

14.2–18.4 **S3** [turn] —two more people from the rotary club are coming. So we should plan for fifteen at least. ▷ yields to S1

18.3–23.1 **S1** [turn] Good to know. Mm, I want to split everyone into small groups so it doesn't get chaotic. ▷ yields to S3

20.4–21.5 **S2** [failed claim] Yeah, that makes sense.

23.4–26.7 **S3** [turn] We need someone to run the check-in table, too, so maybe don't put everyone straight into cleanup. > cut off

25.6–30.9 **S2** [interruption] —Oh, also, I was going to ask about supplies. Do we have enough vests and gloves for fifteen? > takes floor from S3

31.0–34.3 **S2** [turn] I saw a box in the storage room but it didn't look like a full set. > yields to S1

34.3–39.1 **S1** [turn] I can check inventory tomorrow. If we're short, I can pick some up in the morning.

35.5–36.2 **S2** [backchannel] Thanks, Aaron.

38.0–39.2 **S3** [backchannel] Mm, appreciate that.

39.2–45.0 **S1** [turn] No problem. Uh—let's talk group leads next. Do we have volunteers or do we need to assign? > yields to S3

44.9–47.3 **S3** [turn] I can run the playground crew, makes things quicker. > yields to S2

47.2–51.5 **S2** [turn] I can help with—well, maybe the trail section? Not as many heavy items there.

48.8–51.1 **S3** [failed claim] Do you want someone paired with you in case it's muddy?

51.6–53.4 **S2** [turn] —oh, yeah, that's a good idea. Yeah. > yields to S1

53.8–56.1 **S1** [turn] I can see if Susan's free for the trail group—

55.4–57.5 **S3** [failed claim] She usually brings her own gloves, too.

55.6–56.3 **S2** [laughter] Haha, right.

56.2–60.9 **S1** [turn] Okay, so check-in table—anyone want to handle sign-ins or should I just do it? > yields to S3

60.8–64.6 **S3** [turn] I'll swing by early and help set up, but I won't stay at the table all morning. > yields to S2

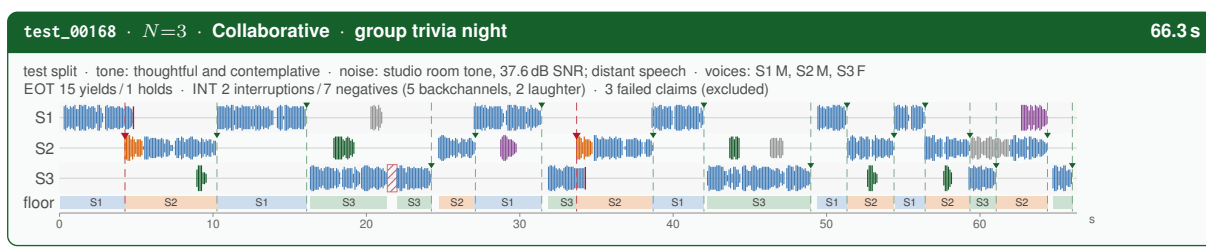
64.5–67.9 **S2** [turn] I can rotate with you for the first half hour or so, if that helps. > yields to S1

68.2–73.5 **S1** [turn] Perfect. I think that's all the big stuff. Any other concerns or last-minute ideas? > yields to S3

73.4–76.5 **S3** [turn] Just hoping the weather holds up. Otherwise, we're good to go. > yields to S1

75.5–76.5 **S2** [backchannel] Yeah, fingers crossed.

76.9–80.7 **S1** [turn] Alright, I'll send out a summary email in an hour. Thanks guys. > yields, end



Transcript of test_00168

start–end (s) speaker [ground-truth label] text > outcome

0.3–4.8 **S1** [turn] Um, so the next trivia question says, uh, 'Which planet is known as the morning star?' > cut off

4.3–5.5 **S2** [interruption] Oh that one's Venus— > takes floor from S1

5.6–10.3 **S2** [turn] Seriously, I'm pretty sure it's Venus. I just, like, read an article on space stuff last week. > yields to S1

9.0–9.7 **S3** [backchannel] Right.

10.3–16.1 **S1** [turn] Yeah, that sounds right, but isn't Mercury also called that sometimes? Or, uh, am I mixing it up? > yields to S3

16.4–19.6 **S3** [turn] No, Venus is the main one, I think. Mercury's too faint.

17.9–19.3 **S2** [backchannel] Yeah, yeah, exactly.

19.7–21.4 **S3** [turn] We locking in Venus then? > holds, pauses

20.3–21.1 **S1** [failed claim] Uh, one sec—

22.0–24.2 **S3** [turn] —Unless you think it's Mercury after all? > yields to S2

24.7–27.1 **S2** [turn] No way, Venus is, like, so much brighter. > yields to S1

27.0–31.4 **S1** [turn] Yeah, okay, let's go with Venus then. Sorry, I just overthink these sometimes. > yields to S3

28.8–29.8 **S2** [laughter] Haha, same.

31.9–34.3 **S3** [turn] Alright, next one: 'Which author wrote The Old Man and the Sea?' Easy. > cut off

33.7–34.8 **S2** [interruption] Ernest Hemingway— > takes floor from S3

34.9–38.7 **S2** [turn] Definitely Hemingway. That novella, right? He won the Nobel or... Pulitzer? > yields to S1

38.6–42.0 **S1** [turn] I think, uh, Pulitzer? The Nobel was for all his work, I think. > yields to S3

42.2–44.4 **S3** [turn] Pulitzer for that book. Confirmed.

43.7–44.4 **S2** [backchannel] Great, locked.

44.5–47.3 **S3** [turn] Next: 'Largest internal organ in humans?'

46.3–47.2 **S2** [failed claim] Ooh, is that—

47.4–49.0 **S3** [turn] —Liver. Has to be liver. > yields to S1

49.4–51.3 **S1** [turn] Wait, isn't skin considered an organ though? > yields to S2

51.4–54.4 **S2** [turn] But skin's external, right? The question said internal. > yields to S1

52.7–53.4 **S3** [backchannel] Right.

54.4–56.4 **S1** [turn] Oh, right! Missed that. Yeah, liver then. > yields to S2

56.4–59.3 **S2** [turn] Liver is actually, like, shockingly heavy for its size. > yields to S3

57.6–58.2 **S3** [backchannel] Huh.

59.3–61.1 **S3** [turn] —It shouldn't be that color either. > yields to S2

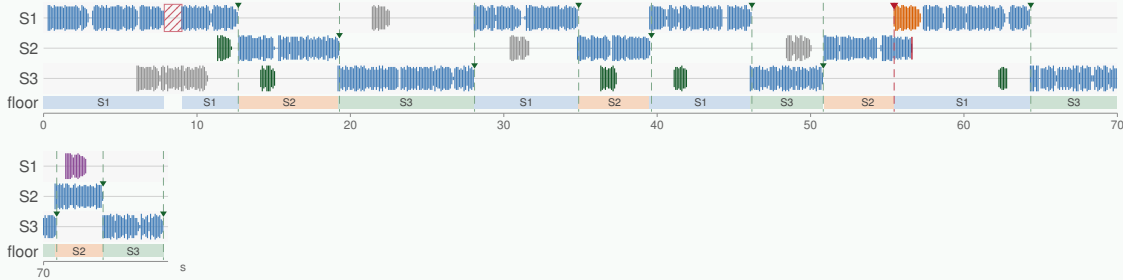
59.4–61.9 **S2** [failed claim] No, seriously! I saw one in biology class—

62.0–64.4 **S2** [turn] Exactly! It was like dark reddish brown. > yields to S3

62.7–64.4 **S1** [laughter] Haha, yeah, kind of gross honestly.

64.8–66.0 **S3** [turn] Ready for the next? > yields, end

test split · tone: curious and inquisitive · noise: quiet apartment, 31.8 dB SNR; dishes clinking, notification chime · voices: S1 M, S2 F, S3 F
EOT 11 yields / 1 holds · INT 1 interruptions / 6 negatives (5 backchannels, 1 laughter) · 4 failed claims (excluded)



Transcript of test_00179

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–7.9 **S1** [turn] Um, so I—I was just thinking about that last play, like, do you think the coach actually meant to run that zone defense or was it a miscommunication? ▷ holds, pauses

6.1–10.7 **S3** [failed claim] Yeah, I was wondering the same! It looked like two players kept calling to switch and then—

9.0–12.7 **S1** [turn] —they both kinda chased the ball, right, so the lane was totally open. ▷ yields to S2

11.3–12.3 **S2** [backchannel] Exactly.

12.7–19.3 **S2** [turn] But you know, I actually rewatched that replay and it seemed like the point guard started to gesture for a switch a split second too late. ▷ yields to S3

14.2–15.1 **S3** [backchannel] Right right.

19.2–23.2 **S3** [turn] It almost makes me wonder if they actually practiced that rotation at all—

21.4–22.6 **S1** [failed claim] Wait, but what if—

23.3–28.1 **S3** [turn] —because their footwork was totally off, especially number twelve. Did you see her stumble? ▷ yields to S1

28.1–34.9 **S1** [turn] Yeah, I saw that, but, um, what if that was just nerves though? Like, maybe she wasn't sure if she was supposed to hedge or switch all the way. ▷ yields to S2

30.4–31.7 **S2** [failed claim] That's a good point.

34.8–39.6 **S2** [turn] Cause the coach was shouting something from the bench, but the cameras like, totally missed what he was signaling. ▷ yields to S1

36.3–37.4 **S3** [backchannel] Oh for sure.

39.5–46.2 **S1** [turn] Did either of you notice, um, like, the opposing center waving at the ref right before the play? Was she complaining about something? ▷ yields to S3

41.1–42.0 **S3** [backchannel] I did!

46.1–50.8 **S3** [turn] Actually, I think she was calling out a push under the basket but nobody paid attention. ▷ yields to S2

48.4–50.1 **S2** [failed claim] She looked pretty frustrated.

50.9–56.6 **S2** [turn] What did you think of, like, the way they kept switching screens in the third quarter? I felt like that's where the momentum shifted. ▷ cut off

55.4–57.2 **S1** [interruption] Sorry, can I just jump in?— ▷ takes floor from S2

57.4–64.4 **S1** [turn] I absolutely agree, like, as soon as they started that, the other team couldn't get any open looks at all. The spacing just collapsed. ▷ yields to S3

62.3–62.9 **S3** [backchannel] Mm-hmm.

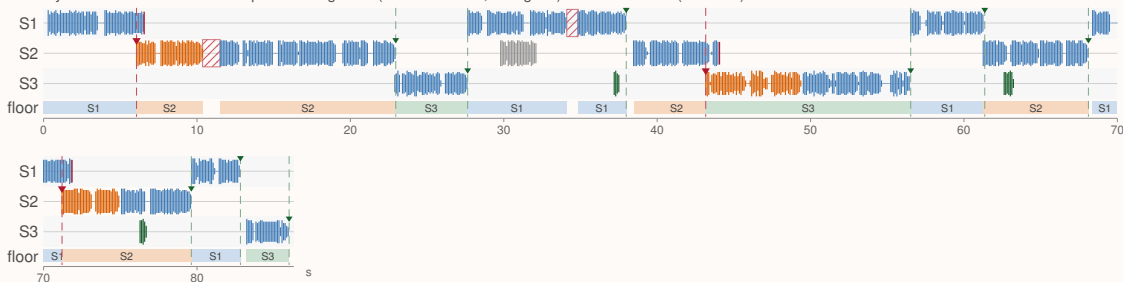
64.3–70.9 **S3** [turn] But then, like, two minutes later, didn't they just go straight back to man-to-man and suddenly gave up three baskets in a row? ▷ yields to S2

70.8–73.9 **S2** [turn] Yes! I kept yelling, why are you changing it when it's working? Haha. ▷ yields to S3

71.4–72.8 **S1** [laughter] Haha, exactly.

73.9–77.8 **S3** [turn] I know, they just overthink it sometimes. It's kind of fascinating. ▷ yields, end

dev split · tone: thoughtful and contemplative · noise: office hum, 22.4 dB SNR; phone ring · voices: S1 F, S2 M, S3 M
EOT 9 yields / 2 holds · INT 3 interruptions / 3 negatives (3 backchannels, 0 laughter) · 1 failed claims (excluded)



Transcript of dev_00018

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.6 **S1** [turn] Okay, the next question is about the capital city of Kazakhstan, and I'm telling you right now it's not Almaty anymore, it's Astana, or, uh, I think they renamed it again—Nur-Sultan, maybe? ▷ cut off

6.1–10.4 **S2** [interruption] I mean, technically, but didn't they change it back to Astana last year? ▷ takes floor from S1 ▷ holds, pauses

11.5–23.0 **S2** [turn] I read a whole article, like, in September or October, they officially re-switched the name from Nur-Sultan to Astana. So if the trivia's up to date, it should be Astana again. ▷ yields to S3

22.9–27.6 **S3** [turn] Sure, but nobody actually calls it Nur-Sultan unless they're reading the news. ▷ yields to S1

27.6–34.1 **S1** [turn] That's not true, I listened to a whole geopolitics podcast, and everybody on there kept saying Nur-Sultan— ▷ holds, pauses

29.8–32.2 **S2** [failed claim] Yeah, but podcasts lag behind, right?

34.8–38.0 **S1** [turn] —because that's what the government officially called it for three years. ▷ yields to S2

37.2–37.6 **S3** [backchannel] hmm

38.5–44.1 **S2** [turn] Okay, but look, if they ask for the current capital, it's definitely Astana now, according to— ▷ cut off

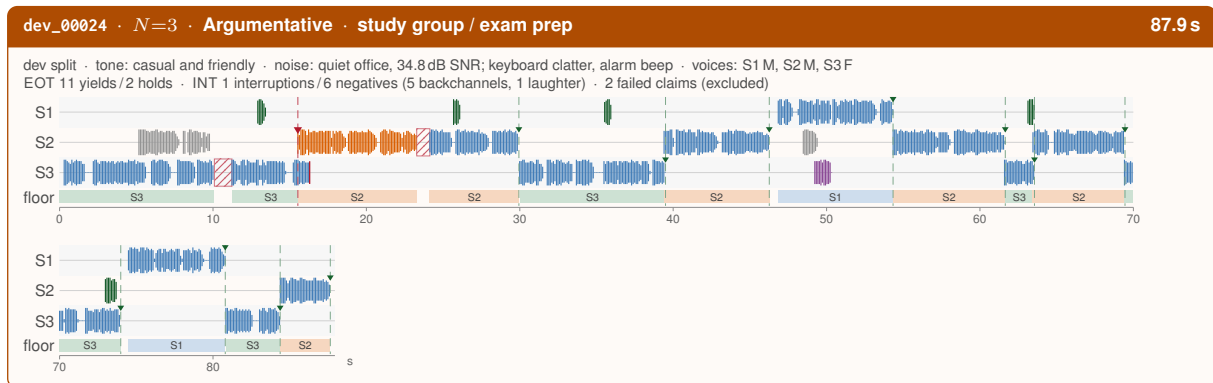
43.2–49.4 **S3** [interruption] Wait, but when did the trivia questions get written? That's what matters. It's always like six months behind. ▷ takes floor from S2

49.5–56.5 **S3** [turn] So if these questions are old, it's probably Nur-Sultan, even if it's not anymore. They don't update fast. ▷ yields to S1

56.5–61.4 **S1** [turn] Exactly. That's why I'm saying, we should just write down both and see what the host says. ▷ yields to S2

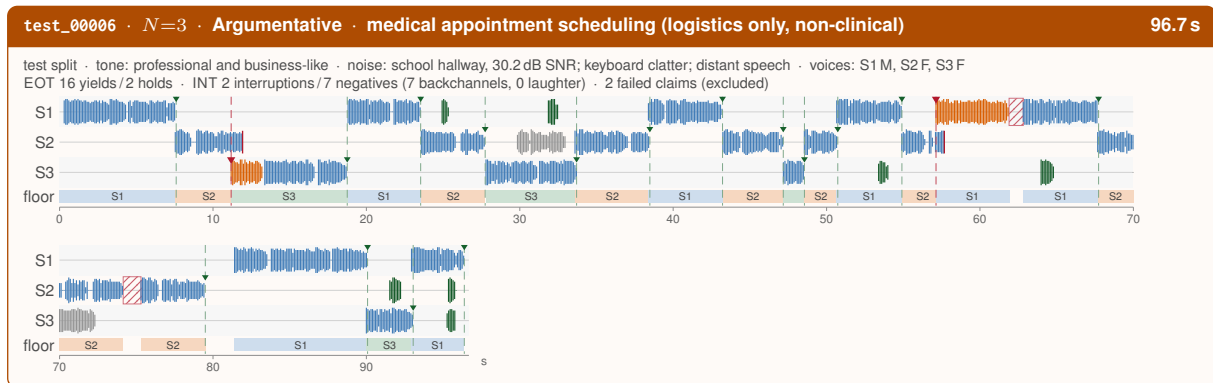
61.3–64.9 **S2** [turn] But we won't get credit for both, they're gonna want the most recent—

62.6–63.2 **S3** [backchannel] no way
 65.0–68.1 **S2** [turn] —one in their answer key. Otherwise it's chaos. ▷ yields to S1
 68.4–71.9 **S1** [turn] Well, do you want to risk it? Because if we get it wrong on a technicality— ▷ cut off
 71.2–75.0 **S2** [interruption] No, I want us to actually win, so I'm writing Astana. ▷ takes floor from S1
 75.1–79.6 **S2** [turn] Final answer, Astana. If we lose, I'll buy the next round, agreed? ▷ yields to S1
 76.3–76.8 **S3** [backchannel] deal
 79.7–82.8 **S1** [turn] Alright, I'm holding you to that, but I still think it's a gamble. ▷ yields to S3
 83.2–86.0 **S3** [turn] At least it's more interesting than the flag questions. ▷ yields, end



Transcript of dev_00024 start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–10.1 **S3** [turn] Alright, let's settle this. The practice exam says mitochondria is the site of glycolysis, but that's flat out wrong. It's in the cytoplasm, everyone knows that. ▷ holds, pauses
 5.2–9.8 **S2** [failed claim] No, wait, but isn't the mitochondria still involved? Like, once pyruvate forms—
 11.2–16.3 **S3** [turn] —yes, but glycolysis itself doesn't go into the mitochondria. It literally finishes before anything crosses the membrane. ▷ cut off
 12.9–13.4 **S1** [backchannel] Right
 15.5–23.3 **S2** [interruption] Okay, but the question says 'where does glucose breakdown happen,' not 'where is ATP generated,' and the mitochondria is, like, the powerhouse— ▷ takes floor from S3 ▷ holds, pauses
 24.1–29.9 **S2** [turn] —so wouldn't it make sense that, like, most of the breakdown process happens there? I just think it's not black and white. ▷ yields to S3
 25.7–26.2 **S1** [backchannel] Uh-huh
 30.0–39.5 **S3** [turn] But the actual initial steps, and that's what glycolysis means, is in the cytoplasm. You can't just call the whole cellular respiration process 'mitochondrial.' ▷ yields to S2
 35.5–36.0 **S1** [backchannel] Sure
 39.4–46.3 **S2** [turn] I mean, look, literally every diagram highlights the mitochondria. Are you telling me all those textbooks are just misleading us? ▷ yields to S1
 46.9–50.3 **S1** [turn] Yeah, probably. Diagrams skip steps all the time.
 48.5–49.4 **S2** [failed claim] No, but—
 49.2–50.3 **S3** [laughter] Haha, exactly
 50.4–54.3 **S1** [turn] —like, there's literally an arrow that goes nowhere, half the time. ▷ yields to S2
 54.3–61.6 **S2** [turn] Alright, fine, but I still think if you wrote 'cytoplasm' as the answer, they'd mark it wrong just because everyone expects 'mitochondria.' ▷ yields to S3
 61.6–63.5 **S3** [turn] That's not how grading works, trust me. ▷ yields to S2
 63.1–63.6 **S1** [backchannel] Yeah
 63.5–69.5 **S2** [turn] I seriously doubt it. My chem teacher last year totally gave me zeros for putting details they didn't expect. ▷ yields to S3
 69.4–74.0 **S3** [turn] That's chemistry, not biology. And if you explain, they actually read your logic. ▷ yields to S1
 73.0–73.8 **S2** [backchannel] No way
 74.5–80.8 **S1** [turn] You could always add a note like 'see page 42,' if you want to cover yourself. Not that it matters. ▷ yields to S3
 80.8–84.4 **S3** [turn] Honestly, that'd just confuse them. Just write 'cytoplasm.' ▷ yields to S2
 84.4–87.6 **S2** [turn] Haha, you guys are gonna get me in trouble on this exam. ▷ yields, end



Transcript of test_00006 start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–7.6 **S1** [turn] Alright, so the Jones appointment—you want to just push it to Friday afternoon? That'd clear our Tuesday, honestly makes my life easier. ▷ yields to S2
 7.5–11.9 **S2** [turn] No, that doesn't work. Friday's already full, plus Dr. Patel isn't even here past two. ▷ cut off
 11.2–13.2 **S3** [interruption] Can we just look at Thursday instead— ▷ takes floor from S2
 13.4–18.8 **S3** [turn] —Thursday morning is open, and that's way better for the Jones family. They've emailed twice about mornings. ▷ yields to S1

18.8–23.5 **S1** [turn] Right, but then you’ve got the pre-op consults stacked in there. That’s gonna be tight timing, isn’t it? ▷ yields to S2

23.6–27.7 **S2** [turn] Yeah, but we can just move Ms. Han to later. She’s local, she’s flexible. ▷ yields to S3

24.9–25.4 **S1** [backchannel] Mm-hm.

27.8–31.2 **S3** [turn] I already shifted Han last week. We can’t just keep bumping her, that’s—

29.8–33.0 **S2** [failed claim] —she said she was fine with anytime, I literally called her—

31.3–33.7 **S3** [turn] —yeah, but she’s gonna notice. It’s not a good look. ▷ yields to S2

31.9–32.5 **S1** [backchannel] Right, right.

33.6–38.5 **S2** [turn] Look, I’m not suggesting we push her hours, just by thirty minutes. I don’t see the big issue. ▷ yields to S1

38.4–43.2 **S1** [turn] Except it ripples. If Han moves, then the labs need to reschedule, then billing complains— ▷ yields to S2

43.2–47.2 **S2** [turn] No, that’s not true. The lab is open the whole morning. I already checked. ▷ yields to S3

47.2–48.6 **S3** [turn] You sent them an email? ▷ yields to S2

48.6–50.7 **S2** [turn] Yes, Jamie confirmed yesterday. ▷ yields to S1

50.7–54.9 **S1** [turn] Okay, but, like, did Jamie actually check with the techs or just say yes to be nice? ▷ yields to S2

53.4–54.0 **S3** [backchannel] Yeah.

54.9–57.7 **S2** [turn] She copied them on the reply. Look, we are stalling here— ▷ cut off

57.1–61.9 **S1** [interruption] Wait, real quick—if we move Jones to Thursday, what’s the backup if the kids get sick again? ▷ takes floor from S2 ▷ holds, pauses

62.8–67.7 **S1** [turn] Because they’ve canceled twice, we can’t keep holding prime slots for them and then have an empty block. ▷ yields to S2

64.0–64.8 **S3** [backchannel] Exactly.

67.7–74.1 **S2** [turn] Then we do overbooking, like we did in February. It’s not perfect, but it worked. We can’t just freeze the schedule for a maybe. ▷ holds, pauses

70.0–72.3 **S3** [failed claim] Overbooking was chaos, I’m not repeating that—

75.3–79.5 **S2** [turn] —well, then you have to pick, either we take the risk or we push them to the waitlist. ▷ yields to S1

81.4–90.1 **S1** [turn] Alright, let’s just—can we get a consensus? Thursday morning, only if Han agrees in writing, and labs confirm? Otherwise, I’m putting Jones on the waitlist. ▷ yields to S3

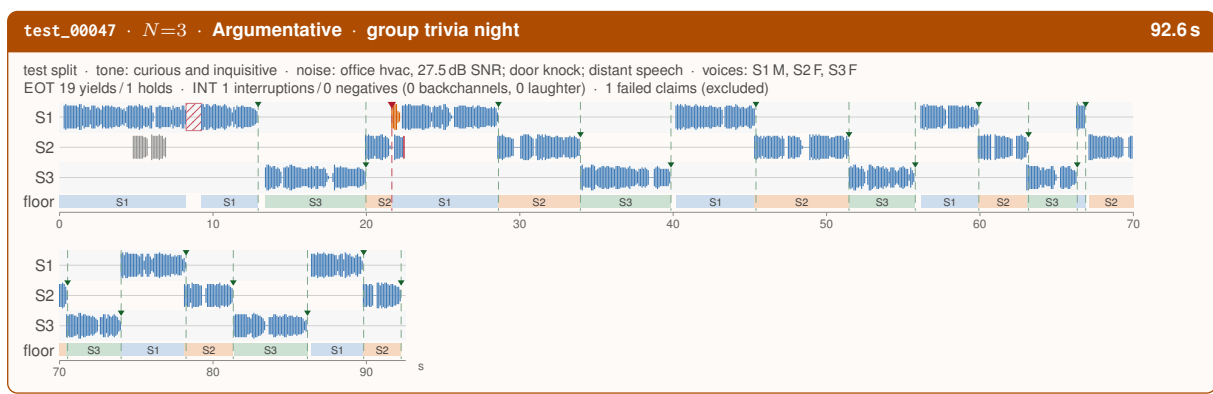
90.0–93.0 **S3** [turn] Fine, but only if Han’s good with it. I’m not fielding complaints. ▷ yields to S1

91.5–92.3 **S2** [backchannel] Agreed.

92.9–96.4 **S1** [turn] Cool, I’ll call Han now. If she pushes back, Jones gets bumped. Deal? ▷ yields, end

95.3–95.9 **S3** [backchannel] Deal.

95.4–95.9 **S2** [backchannel] Done.



Transcript of test_00047

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.2 **S1** [turn] Alright, so the next question is, uh, which planet has the most moons? I’m pretty sure it’s Saturn now, right? They passed Jupiter a couple years back. ▷ holds, pauses

4.8–7.0 **S2** [failed claim] Wait-no, hang on, I just saw an article—

9.3–12.9 **S1** [turn] —I swear, it was Saturn like last year, they found a ton more little ones. ▷ yields to S3

13.4–20.0 **S3** [turn] No, it’s Jupiter. They just updated it again, and Jupiter’s back on top. That was literally in the news like last month. ▷ yields to S2

19.9–22.4 **S2** [turn] Exactly! That’s what I saw. Thank you, Molly. ▷ cut off

21.7–22.2 **S1** [interruption] Hold up— ▷ takes floor from S2

22.3–28.6 **S1** [turn] I’m not buying it. There’s no way they keep flipping it that fast. Are you sure it wasn’t just, like, speculation or something? ▷ yields to S2

28.6–34.0 **S2** [turn] No, I’m serious! Jupiter got, like, twelve surprise moons or something completely ridiculous. ▷ yields to S3

34.0–39.9 **S3** [turn] Yes. It was twelve, and they’re all tiny, irregular rocks basically, but they count. Jupiter’s in the lead. ▷ yields to S1

40.2–45.4 **S1** [turn] Wow, that’s wild. Okay, so are we sure about this, though? I feel like Saturn’s number was massive. ▷ yields to S2

45.3–51.5 **S2** [turn] Yeah, but Jupiter’s is literally, like, eighty-something now. I saw that headline and the picture of all the dots. ▷ yields to S3

51.5–55.8 **S3** [turn] Eighty-seven, I think. Saturn’s trailing by maybe three now. Jupiter leapfrogged. ▷ yields to S1

56.2–59.9 **S1** [turn] Alright, fine, I’ll trust you two. But if we’re wrong, you’re both buying drinks. ▷ yields to S2

59.9–63.2 **S2** [turn] Deal. But we’re not. I’d stake, like, at least...a soda on it. ▷ yields to S3

63.1–66.3 **S3** [turn] Heh, you’ll owe us drinks instead. Ready for the next one? ▷ yields to S1

66.3–66.9 **S1** [turn] Yeah, shoot. ▷ yields to S2

67.1–70.5 **S2** [turn] Alright, who painted the ceiling of the Sistine Chapel? Easy, right? ▷ yields to S3

70.5–74.0 **S3** [turn] Michelangelo. That’s not even a debate. Can we get a harder question? ▷ yields to S1

74.0–78.2 **S1** [turn] Wait, wasn’t Leonardo involved too, or am I just mixing up Ninja Turtles? ▷ yields to S2

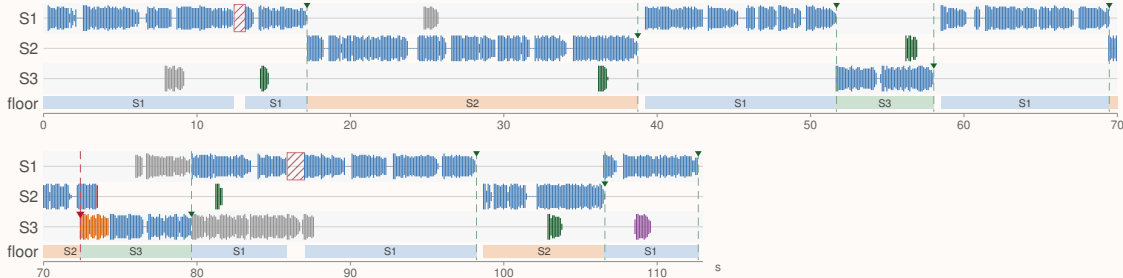
78.1–81.3 **S2** [turn] Haha, oh my god, Ninja Turtles logic strikes again. ▷ yields to S3

81.4–86.2 **S3** [turn] No, just Michelangelo. Not Da Vinci. He was painting mysterious smiles somewhere else. ▷ yields to S1

86.4–89.8 **S1** [turn] Fair, fair. Okay, I’m done arguing. For like, five seconds. Next! ▷ yields to S2

89.8–92.3 **S2** [turn] Haha, sure. Let’s try geography this time. ▷ yields, end

test split · tone: professional and business-like · noise: office hvac, 32.5 dB SNR; door knock, phone ring · voices: S1 F, S2 M, S3 F
EOT 9 yields/2 holds · INT 1 interruptions/6 negatives (5 backchannels, 1 laughter) · 5 failed claims (excluded)



Transcript of test_00061

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–12.4 **S1** [turn] Let's just start with the basic fact: the policy as drafted is unworkable for small businesses. We've seen it time and again, these blanket regulations end up hurting the people they're supposed to help. ▷ holds, pauses

7.9–9.2 **S3** [failed claim] I think that's a little—

13.2–17.2 **S1** [turn] —hurting them, because compliance costs skyrocket and resources dry up. ▷ yields to S2

14.1–14.7 **S3** [backchannel] Right

17.2–29.5 **S2** [turn] But if you look at the numbers, there's evidence that these kinds of policies actually improve employee retention and productivity. Larger companies absorb it, yes, but there are grants tied into the rollout for smaller ones.

24.8–25.8 **S1** [failed claim] That's theoretical—

29.6–38.7 **S2** [turn] —and the pilot cities showed a net positive, especially in hospitality sectors. I just don't think we can dismiss it as unworkable without acknowledging that context. ▷ yields to S1

36.2–36.8 **S3** [backchannel] Uh-huh

39.2–51.7 **S1** [turn] But the context is different for a bakery with six employees versus a corporation with six thousand. I mean, for some of these shops, the margin is so thin, anything extra could shut them down. ▷ yields to S3

51.7–58.0 **S3** [turn] But isn't that what the sliding scale is supposed to address? There's exemptions up to a certain revenue threshold, if I'm not mistaken. ▷ yields to S1

56.2–57.0 **S2** [backchannel] Exactly

58.5–69.5 **S1** [turn] Yet enforcement is patchy. Look, the local coffee shop owner from last week's council hearing said she'd have to cut hours to stay open, even with the supposed exemptions. ▷ yields to S2

69.4–73.5 **S2** [turn] I just don't think anecdotal evidence is enough. We need to focus on aggregate outcomes, not single cases. ▷ cut off

72.4–74.3 **S3** [interruption] But personal stories matter, too— ▷ takes floor from S2

74.4–79.6 **S3** [turn] Because if we ignore what's happening on the ground, policies become completely disconnected from reality. ▷ yields to S1

76.0–79.6 **S1** [failed claim] Yes, and the rollout last quarter caught a lot of people off guard—

79.7–85.9 **S1** [turn] —it's just that nobody outside the business district even heard about them. Outreach isn't just checking a box. ▷ holds, pauses

79.7–83.4 **S3** [failed claim] —but didn't they hold three town halls? The communication was there, it's just—

81.2–81.7 **S2** [backchannel] Right

83.5–87.7 **S3** [failed claim] But do you really think more outreach would've solved the compliance costs? Or is it more—

87.0–98.2 **S1** [turn] —well, it wouldn't fix everything but it's a start. Accountability on both sides matters. If the city made consultation more than just a formality, maybe we'd be looking at something more workable. ▷ yields to S2

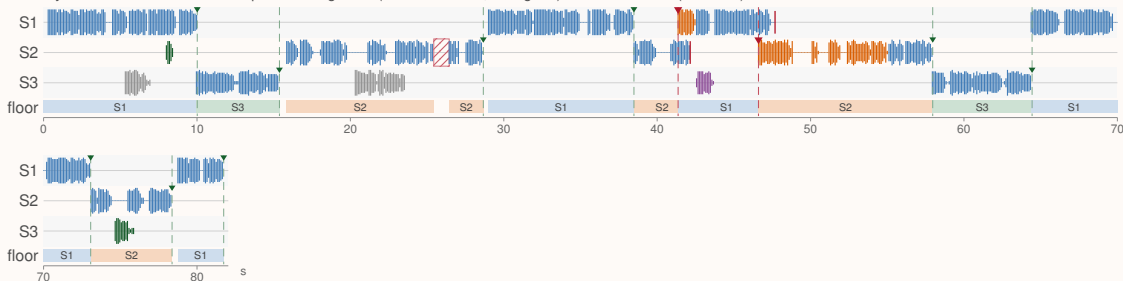
98.7–106.6 **S2** [turn] Fair point. But to move forward, compromise is key. Businesses and lawmakers both need to get out of their echo chambers or nothing gets done. ▷ yields to S1

102.9–103.8 **S3** [backchannel] Absolutely

106.5–112.7 **S1** [turn] Couldn't agree more, but let's make sure compromise isn't just code for watering things down until they're meaningless. ▷ yields, end

108.6–109.6 **S3** [laughter] Haha, right

test split · tone: curious and inquisitive · noise: school hallway, 27.9 dB SNR; notification chime; distant speech · voices: S1 M, S2 F, S3 M
EOT 9 yields/1 holds · INT 2 interruptions/3 negatives (2 backchannels, 1 laughter) · 2 failed claims (excluded)



Transcript of test_00067

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.8 **S1** [turn] Alright, the next question is, uh, which planet spins the fastest, right? I think it's Jupiter. It's got the shortest day.

5.3–7.0 **S3** [failed claim] No no, Saturn—

6.9–10.0 **S1** [turn] —at least that's what I remember from, like, third grade science. ▷ yields to S3

8.0–8.5 **S2** [backchannel] mm-hm

10.0–15.4 **S3** [turn] But Saturn's day is, what, ten hours or something? I swear it's Saturn that's the fastest, not Jupiter. ▷ yields to S2

15.8–25.4 **S2** [turn] Are you both just guessing because you like big planets? I feel like this is one of those weird Mercury things. ▷ holds, pauses

20.3–23.6 **S3** [failed claim] No, no, Mercury is slow, it's got that crazy long year—

26.4–28.7 **S2** [turn] —oh, right, the whole day-night thing. ▷ yields to S1

29.0–38.5 **S1** [turn] Actually, Saturn and Jupiter both spin super fast, but I remember something about Jupiter being fastest overall. Like, it's under ten hours. Saturn's a little

slower. ▷ yields to S2

38.5–42.2 **S2** [turn] Well, if we get it wrong, at least we'll have argued about it for half the round. ▷ cut off

41.4–42.5 **S1** [interruption] But wait, hold on— ▷ takes floor from S2

42.6–43.7 **S3** [laughter] Haha, right

42.6–47.7 **S1** [turn] I just remembered—Jupiter's got those huge storms because of how fast it spins, right? Isn't that the main thing? ▷ cut off

46.6–55.0 **S2** [interruption] But doesn't Saturn have more wind speed? Like, isn't that why its rings are so weirdly sharp? ▷ takes floor from S1

55.1–58.0 **S2** [turn] I think you're just making up storm facts now. ▷ yields to S3

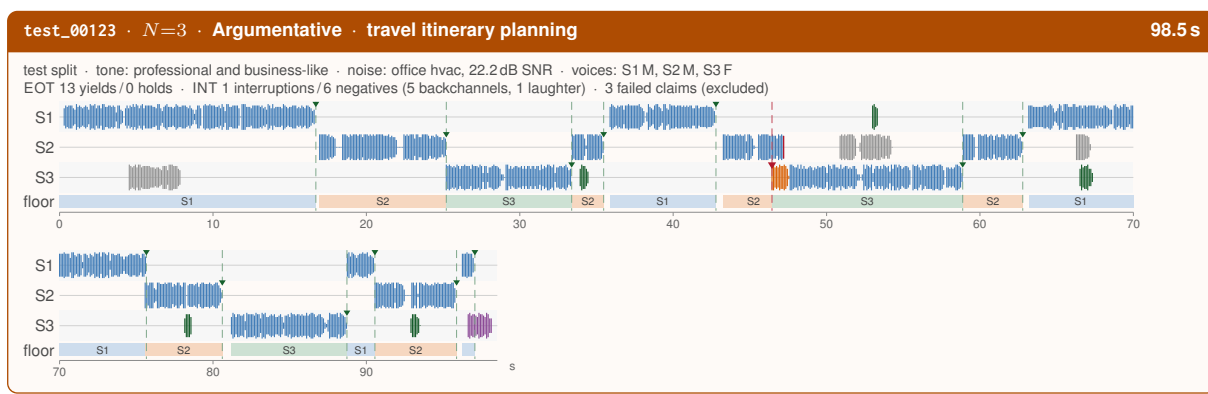
57.9–64.4 **S3** [turn] No, she's right. Saturn's winds are insane, but that doesn't mean it spins faster. That's just upper atmosphere, isn't it? ▷ yields to S1

64.4–73.1 **S1** [turn] Exactly, it's about the core rotation. Jupiter's day is, like, 9 hours and 56 minutes. Saturn's is just a hair longer, 10 and some change. ▷ yields to S2

73.1–78.4 **S2** [turn] Fine, Jupiter then. But if we're wrong, I'm blaming Wikipedia. ▷ yields to S1

74.7–75.9 **S3** [backchannel] Yeah, agreed.

78.8–81.7 **S1** [turn] Okay, locked in: Jupiter. Next question, please. ▷ yields, end



Transcript of test_00123

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.3 **S1** [turn] So, um, I think we should prioritize getting to Berlin first, since the conference starts the morning of the seventh, and that means flying in on the sixth, right?

4.5–7.9 **S3** [failed claim] But if we go straight to Berlin, we miss seeing Prague, and that was—

9.4–16.7 **S1** [turn] —uh, I know we discussed Prague before, but I really think the logistics get tricky if we try to squeeze it in before the conference. ▷ yields to S2

17.0–25.2 **S2** [turn] Hang on, both of you. Why not just do Berlin, then Prague, and then Vienna after the conference? That makes the most sense geographically and timing-wise. ▷ yields to S3

25.2–33.4 **S3** [turn] But, uh, the train from Berlin to Prague is like, what, four hours? That cuts into almost a whole day post-conference if we only have two days left. ▷ yields to S2

33.4–35.5 **S2** [turn] It's three and a half. I checked last night. ▷ yields to S1

33.9–34.5 **S3** [backchannel] Right

35.9–42.8 **S1** [turn] But that's still like, a lot of moving around? Plus, um, flights back from Vienna are way more expensive than from Berlin. ▷ yields to S2

43.3–47.2 **S2** [turn] Actually, if we fly home from Prague, the fare's about the same as Berlin, I already checked aggregators. ▷ cut off

46.4–47.6 **S3** [interruption] Okay, but I— ▷ takes floor from S2

47.6–55.5 **S3** [turn] I really just don't think packing three cities into five days is going to feel useful. It's just, like, constant transit and no down time.

50.9–54.3 **S2** [failed claim] We can cut Vienna then, focus on Berlin and Prague, and it would—

53.0–53.4 **S1** [backchannel] mm-hm

55.6–58.9 **S3** [turn] —but you both said seeing Vienna was a must, two calls ago! ▷ yields to S2

58.9–62.8 **S2** [turn] I know, I know, but if we can't do all three, we have to compromise somewhere. ▷ yields to S1

63.2–68.7 **S1** [turn] Or maybe we do Vienna first, skip Prague, and spend real time in Berlin after the conference?

66.3–67.2 **S2** [failed claim] No, look, if we—

66.5–67.4 **S3** [backchannel] But that's—

68.8–75.7 **S1** [turn] —honestly, I'm just worried we're overcomplicating things and setting ourselves up for disappointment if we rush every stop. ▷ yields to S2

75.6–80.6 **S2** [turn] That's fair, but I still think we can make two cities work if we just commit to dropping one entirely. ▷ yields to S3

78.2–78.6 **S3** [backchannel] yeah

81.2–88.7 **S3** [turn] So...what if we book Berlin and Prague, and if the timing looks tight, we just add a Vienna day trip at the last minute? Can we agree on that? ▷ yields to S1

88.8–90.6 **S1** [turn] That sounds... reasonable, actually. ▷ yields to S2

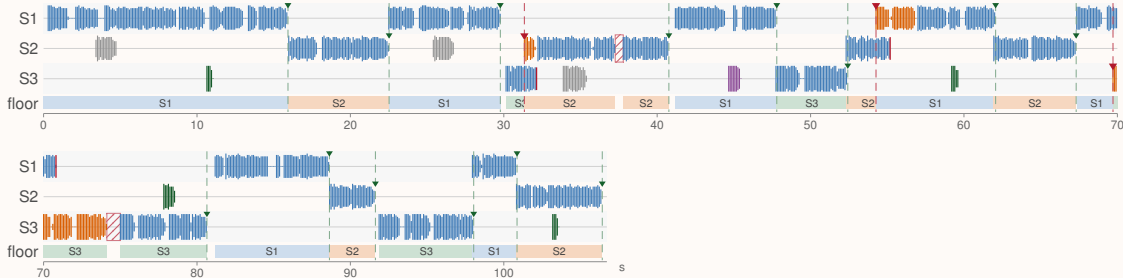
90.6–95.9 **S2** [turn] Alright, let's tentatively lock that in then. I'll check train schedules and send options after this call. ▷ yields to S1

92.9–93.5 **S3** [backchannel] great

96.3–97.1 **S1** [turn] Thanks, everyone. ▷ yields, end

96.6–98.2 **S3** [laughter] haha, we survived it.

test split · tone: thoughtful and contemplative · noise: studio room tone, 28.8 dB SNR; keyboard clatter · voices: S1 F, S2 F, S3 M
EOT 14 yields / 2 holds · INT 3 interruptions / 5 negatives (4 backchannels, 1 laughter) · 3 failed claims (excluded)

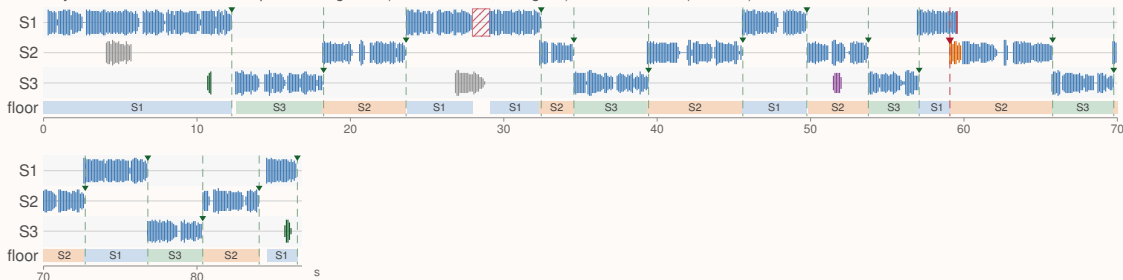


Transcript of test_00154

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.7 **S1** [turn] Alright, I'm just going to say it—remote work is here to stay, and honestly, most companies pushing people back to the office are just stuck in the past.
3.4–4.8 **S2** [failed claim] That's not—can I just—
8.9–15.9 **S1** [turn] —there's no productivity loss! If anything, people get more done at home. I mean, look at every survey from the last year. ▷ yields to S2
10.6–11.1 **S3** [backchannel] mm-hm
16.0–22.5 **S2** [turn] But productivity isn't everything. You're ignoring collaboration and company culture, which don't magically happen over Zoom. ▷ yields to S1
22.5–26.5 **S1** [turn] But Slack exists for a reason! You can have real-time conversations—
25.4–26.8 **S2** [failed claim] That's not the same as—
26.6–29.8 **S1** [turn] —and team-building doesn't have to mean forced happy hours either. ▷ yields to S3
30.2–32.1 **S3** [turn] I think both sides are overstated a bit. ▷ cut off
31.3–32.1 **S2** [interruption] How so? ▷ takes floor from S3
32.2–37.3 **S2** [turn] I mean, there's got to be a line where you draw—if everyone's disconnected, it changes the work itself. ▷ holds, pauses
33.9–35.4 **S3** [failed claim] No, I agree there's a line, but—
37.8–40.8 **S2** [turn] —the energy, the ideas, those little chats... they disappear. ▷ yields to S1
41.2–47.8 **S1** [turn] But can't we just admit those chats are distractions half the time? Like, not everyone wants to bond at the printer, you know? ▷ yields to S3
44.7–45.5 **S3** [laughter] haha, true
47.7–52.4 **S3** [turn] I do think there's something lost, but maybe we're idealizing what office life used to be. ▷ yields to S2
52.3–55.2 **S2** [turn] Except it worked. Companies functioned, teams innovated together— ▷ cut off
54.3–56.9 **S1** [interruption] But at what cost? Seriously, the commute alone— ▷ takes floor from S2
57.0–62.1 **S1** [turn] —is hours of lost time every week, plus stress. People cherish work-life balance now. ▷ yields to S2
59.2–59.7 **S3** [backchannel] Right
61.9–67.3 **S2** [turn] But what about junior employees who need mentoring? You know how tough it is to learn just by watching over Zoom? ▷ yields to S1
67.3–70.8 **S1** [turn] Mentoring can happen remotely! There are entire programs for it— ▷ cut off
69.7–74.1 **S3** [interruption] But the nuance—the unplanned moments—they're really hard to replace virtually. ▷ takes floor from S1 ▷ holds, pauses
75.0–80.6 **S3** [turn] I'm not anti-remote, just saying some things get lost, and I don't think we've figured out how to fill that gap yet. ▷ yields to S1
77.8–78.6 **S2** [backchannel] Exactly
81.2–88.6 **S1** [turn] I don't know, I think we're underestimating how adaptable people are. Like, the world changed and we found ways to keep connecting. ▷ yields to S2
88.6–91.6 **S2** [turn] But not always better ways. That's the part I can't let go of. ▷ yields to S3
91.9–98.0 **S3** [turn] Maybe it comes down to choice. If people thrive remote, let them. If teams need office time, support that too. ▷ yields to S1
98.0–100.9 **S1** [turn] Yeah, like, flexibility actually means something for once! ▷ yields to S2
100.8–106.4 **S2** [turn] I can agree with that—choice matters, but we need to be honest about what's actually lost when the office goes away. ▷ yields, end
103.2–103.6 **S3** [backchannel] mm-hm

test split · tone: casual and friendly · noise: cafe, 21.7 dB SNR; distant speech · voices: S1 M, S2 M, S3 M
EOT 17 yields / 1 holds · INT 1 interruptions / 3 negatives (2 backchannels, 1 laughter) · 2 failed claims (excluded)

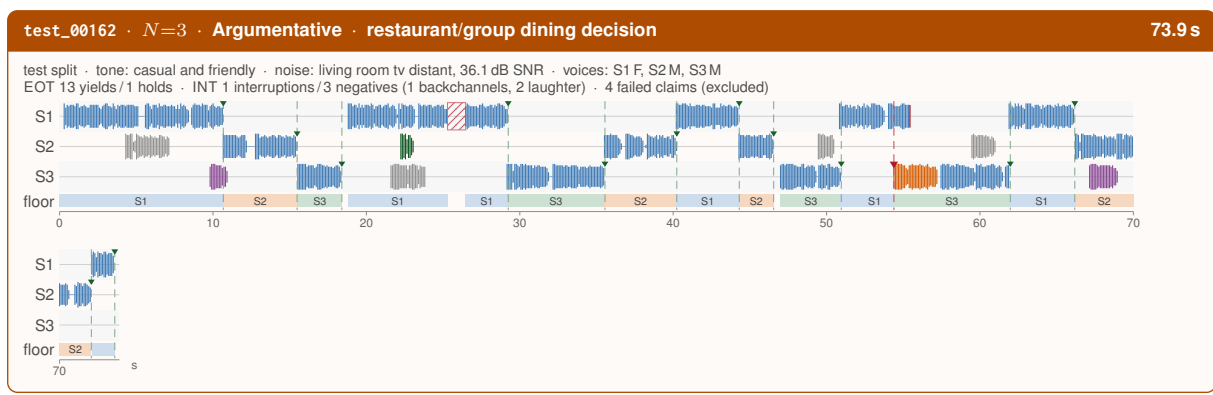


Transcript of test_00155

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–7.9 **S1** [turn] I just think, if we're talking about groceries, it makes more sense to have one person handle the shopping each week instead of splitting everything up.
4.1–5.8 **S2** [failed claim] But then what if someone forgets—
8.0–12.3 **S1** [turn] —something? Yeah, but I mean, we just make a list and, you know, whoever's up for that week follows it. ▷ yields to S3
10.7–11.0 **S3** [backchannel] mm
12.5–18.3 **S3** [turn] Yeah but you two always eat, like, that organic yogurt junk and I don't wanna pay for that stuff every month, man. ▷ yields to S2
18.2–23.6 **S2** [turn] No, we said before you don't have to, right? Like, if you wanna buy your own stuff, just do it. ▷ yields to S1
23.7–28.0 **S1** [turn] Well, but if we start splitting things, then what's even the point of shopping together anymore? ▷ holds, pauses
26.9–28.8 **S3** [failed claim] Because it's still cheaper for basics—
29.1–32.4 **S1** [turn] —yeah like toilet paper, whatever, but it gets complicated fast. ▷ yields to S2
32.3–34.6 **S2** [turn] I dunno, maybe we're overthinking this. ▷ yields to S3

34.6–39.4 **S3** [turn] No, we're not, last time I covered for both of you and no one paid me back for the detergent. ▷ yields to S2
 39.4–45.6 **S2** [turn] Okay, okay, so let's make a spreadsheet? Like, just put up a Google Sheet, write down what everyone buys. ▷ yields to S1
 45.6–49.8 **S1** [turn] But is anyone actually gonna use a spreadsheet? You say that every semester. ▷ yields to S2
 49.8–53.8 **S2** [turn] Hey, last spring I used it. Uh, remember? For, like, almost a month. ▷ yields to S3
 51.5–52.1 **S3** [laughter] haha
 53.8–57.1 **S3** [turn] Bro, nobody filled in anything except you. That's what happened. ▷ yields to S1
 57.0–59.5 **S1** [turn] Exactly! That's why I'm saying, we just assign weeks and move on. ▷ cut off
 59.1–59.9 **S2** [interruption] Okay but— ▷ takes floor from S1
 59.9–65.8 **S2** [turn] What if someone can't afford it on their week? Like, what if I'm broke that time, do I just go hungry? ▷ yields to S3
 65.7–69.8 **S3** [turn] You say that, but you always Venmo me half a week late anyway, so what's the difference? ▷ yields to S2
 69.7–72.7 **S2** [turn] Yeah but that's 'cause you never text me what I owe, man. ▷ yields to S1
 72.6–76.8 **S1** [turn] Can we maybe just make a rule, like, pay within two days of the receipt, no matter who buys? ▷ yields to S3
 76.8–80.4 **S3** [turn] Sure, as long as there's no spreadsheet. I'm not dealing with that again. ▷ yields to S2
 80.4–84.1 **S2** [turn] Deal. And someone take a picture of the receipt or it doesn't count. ▷ yields to S1
 84.5–86.5 **S1** [turn] Fine. I'm in, as long as it's simple. ▷ yields, end
 85.7–86.2 **S3** [backchannel] mm-hm

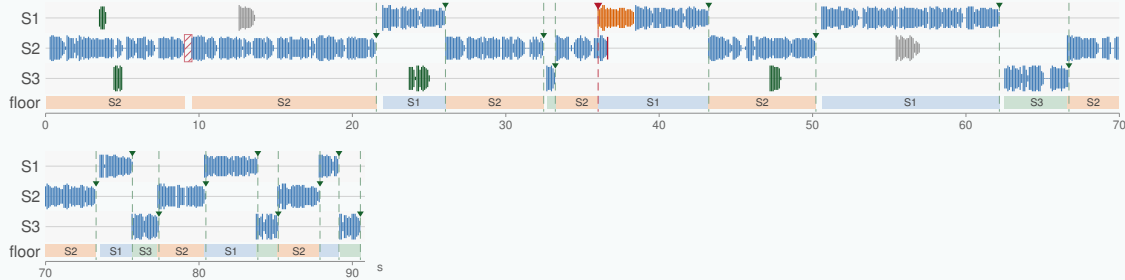


Transcript of test_00162

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.5 **S1** [turn] Okay, so are we actually going to decide on a place or are we just gonna circle around for another hour? Because I swear, every time we do this it's just endless—
 4.3–7.2 **S2** [failed claim] Hang on, let's maybe make a pros and cons list—
 8.6–10.7 **S1** [turn] —debating. I just want food, honestly. ▷ yields to S2
 9.8–11.0 **S3** [laughter] Haha, yeah, seriously.
 10.7–15.5 **S2** [turn] Okay, well, let's make this quick. Thai, pizza, or that new taco place on Elm? ▷ yields to S3
 15.5–18.4 **S3** [turn] Taco place, for sure. We always do pizza, it's boring. ▷ yields to S1
 18.8–25.3 **S1** [turn] Nope, I can't do that taco place, I had food poisoning there last time. Like, violently ill. There's zero chance I'm eating there again. ▷ holds, pauses
 21.6–23.9 **S3** [failed claim] Dude, that was months ago. They probably fixed it.
 22.2–23.1 **S2** [backchannel] Oh, that bad?
 26.5–29.2 **S1** [turn] —No, it was two weeks ago, actually. Not taking that risk. ▷ yields to S3
 29.2–35.6 **S3** [turn] Alright, fine, but I'm not eating pizza again either. Literally every time we hang out it's pizza, and I'm so sick of it. ▷ yields to S2
 35.5–40.2 **S2** [turn] So, what about Thai then? That's still on the table. I mean, Pad Thai, green curry— ▷ yields to S1
 40.2–44.3 **S1** [turn] Yeah, except last time you complained the curry was too spicy and didn't eat half your meal. ▷ yields to S2
 44.3–46.6 **S2** [turn] Okay, I'll just ask for mild, it's fine. ▷ yields to S3
 47.0–51.0 **S3** [turn] Honestly, I'm not that into Thai tonight either. I kinda want tacos, man... ▷ yields to S1
 49.5–50.5 **S2** [failed claim] You're killing me, man.
 50.8–55.4 **S1** [turn] Look, if you want tacos, go eat tacos, but I'm not coming. Like, honestly, pick something we can all actually eat? ▷ cut off
 54.4–57.3 **S3** [interruption] What about burgers? There's that one spot two blocks from here— ▷ takes floor from S1
 57.5–62.0 **S3** [turn] You know, the place with the giant onion rings? That place is never a bad idea. Right? ▷ yields to S1
 59.5–61.0 **S2** [failed claim] Yeah, yeah, I like that idea.
 61.9–66.2 **S1** [turn] Burgers are fine. As long as nobody suddenly remembers they're vegetarian again. ▷ yields to S2
 66.2–68.9 **S2** [turn] Hey, that was one time and it was a phase, okay?
 67.1–69.0 **S3** [laughter] Haha, yeah, 'til you saw the bacon menu.
 68.9–72.1 **S2** [turn] Alright, burger place it is then? We actually done here? ▷ yields to S1
 72.1–73.6 **S1** [turn] Yes, finally. Thank god. ▷ yields, end

dev split · tone: casual and friendly · noise: outdoor park, 33.7 dB SNR · voices: S1 F, S2 M, S3 M
EOT 17 yields / 1 holds · INT 1 interruptions / 4 negatives (4 backchannels, 0 laughter) · 2 failed claims (excluded)



Transcript of dev_00019

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.1 **S2** [turn] Alright, so listen, last night I tried to have a real conversation with Jamie about, you know, why she keeps bringing up stuff from, like, two years ago whenever we're arguing— ▷ holds, pauses

3.5–4.0 **S1** [backchannel] Right.

4.5–5.1 **S3** [backchannel] Classic.

9.5–17.6 **S2** [turn] And as soon as I brought it up, she was like, 'Oh, so now you want to talk about feelings?' Just super sarcastic. I mean, am I crazy or is that just not fair?

12.6–13.6 **S1** [failed claim] So, hang on—

17.7–21.6 **S2** [turn] —I know I've been bad about it before, but I'm actually trying now, you know? ▷ yields to S1

22.0–26.1 **S1** [turn] But were you actually, like, owning up to old stuff or were you just hoping she'd drop it? ▷ yields to S2

23.7–25.0 **S3** [backchannel] Yeah, that matters.

26.1–32.5 **S2** [turn] No, I mean, I said look, I know I used to just shut down, but I want to, like, actually talk through stuff now. I thought that would help. ▷ yields to S3

32.7–33.2 **S3** [turn] Did it? ▷ yields to S2

33.3–36.6 **S2** [turn] Honestly? Nope. She just got defensive and then suddenly it was about how I never listen. ▷ cut off

36.0–38.4 **S1** [interruption] But do you actually listen though? I mean, come on. ▷ takes floor from S2

38.5–43.2 **S1** [turn] Like, when she's talking, are you just waiting for your chance to talk or are you actually hearing her? ▷ yields to S2

43.2–50.2 **S2** [turn] You sound like my therapist, honestly. But yeah, most of the time I'm just internally panicking and trying to find a way to fix it. ▷ yields to S1

47.2–47.9 **S3** [backchannel] Relatable.

50.6–57.5 **S1** [turn] No, but seriously, that's kind of the whole thing. If you can't even stand to sit with her being upset, then it's always gonna turn into a fight.

55.5–57.0 **S2** [failed claim] Okay, philosopher—

57.7–62.2 **S1** [turn] —I'm just saying, you can't shortcut that part. Like, my ex used to do that and it drove me nuts. ▷ yields to S3

62.5–66.7 **S3** [turn] You and Jamie should both just get a hobby. Preferably separately. ▷ yields to S2

66.6–73.3 **S2** [turn] Ha, yeah, maybe I should just get really into, I dunno, bird watching or something and stop overthinking every conversation. ▷ yields to S1

73.5–75.7 **S1** [turn] You're not gonna survive in the woods, man. ▷ yields to S3

75.6–77.4 **S3** [turn] He'd get lost in the parking lot. ▷ yields to S2

77.3–80.5 **S2** [turn] Wow, thanks for the support, guys. I feel totally heard. ▷ yields to S1

80.4–83.8 **S1** [turn] Anytime. We'll be here to roast you after every failed heart-to-heart. ▷ yields to S3

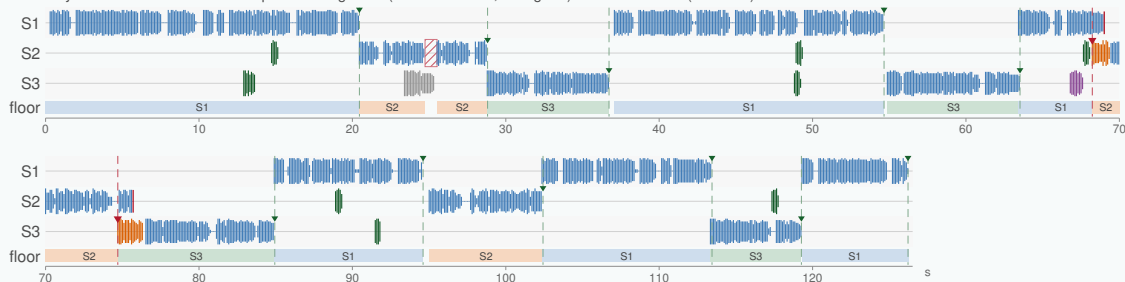
83.7–85.2 **S3** [turn] That's what friends are for. ▷ yields to S2

85.1–87.9 **S2** [turn] Incredible. Maybe next time you can mediate. ▷ yields to S1

87.8–89.1 **S1** [turn] Only if I get snacks. ▷ yields to S3

89.2–90.5 **S3** [turn] Popcorn and chaos. ▷ yields, end

dev split · tone: professional and business-like · noise: office hum, 28.0 dB SNR; distant speech · voices: S1 M, S2 M, S3 F
EOT 11 yields / 1 holds · INT 2 interruptions / 9 negatives (8 backchannels, 1 laughter) · 1 failed claims (excluded)



Transcript of dev_00040

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–20.5 **S1** [turn] Alright, so to kick us off, I wanted to mention that recent survey comparing remote work satisfaction across different industries—it's actually kind of fascinating. They found, um, financial services workers reported the highest productivity gains, while folks in hospitality and retail barely saw any improvement. ▷ yields to S2

12.9–13.7 **S3** [backchannel] interesting

14.7–15.2 **S2** [backchannel] hm

20.5–24.7 **S2** [turn] I guess that tracks. I mean, you can't exactly bartend from Zoom. ▷ holds, pauses

23.4–25.3 **S3** [failed claim] Yeah, but doesn't that kind of ignore—

25.5–28.8 **S2** [turn] —like, the job just doesn't translate. But, go ahead. ▷ yields to S3

28.8–36.7 **S3** [turn] No, I was just gonna say, isn't that a little simplistic? Because even in tech, I know plenty of people who feel less connected to their teams working from home. ▷ yields to S1

37.1–54.7 **S1** [turn] That's definitely fair. Actually, a friend of mine at a software company said the same—her productivity numbers went up at first, but after about six months, she got kind of, uh, isolated. She said group chats can't really replace, like, spontaneous office conversations. ▷ yields to S3

48.8–49.2 **S3** [backchannel] Right

48.9–49.4 **S2** [backchannel] yeah

54.9–63.5 **S3** [turn] I've had that too. My team actually tried this thing where we'd just run a permanent video call like a virtual break room, but honestly, hardly anyone actually used it. > yields to S1

63.4–69.0 **S1** [turn] Haha, I've heard of teams doing that. It's like, you can't force those casual run-ins, right? > cut off

66.8–67.6 **S3** [laughter] haha yeah

67.7–68.2 **S2** [backchannel] nope

68.2–69.4 **S2** [interruption] Actually, on that— > takes floor from S1

69.5–75.7 **S2** [turn] My old boss would, uh, just randomly call people for, like, five-minute check-ins. At first it felt invasive, but some folks liked it. I didn't, but, you know, I'm not exactly— > cut off

74.7–76.4 **S3** [interruption] You're not a small talker, yeah. > takes floor from S2

76.5–84.9 **S3** [turn] But that actually brings up the point about management styles changing with remote work. Some people thrive on that autonomy and others, like, really need structure. > yields to S1

84.9–94.6 **S1** [turn] Definitely. And, the data backs that up—teams with regular but non-intrusive check-ins seem to report higher satisfaction. It's kind of a balancing act. > yields to S2

88.9–89.4 **S2** [backchannel] Sure

91.5–91.9 **S3** [backchannel] mm-hm

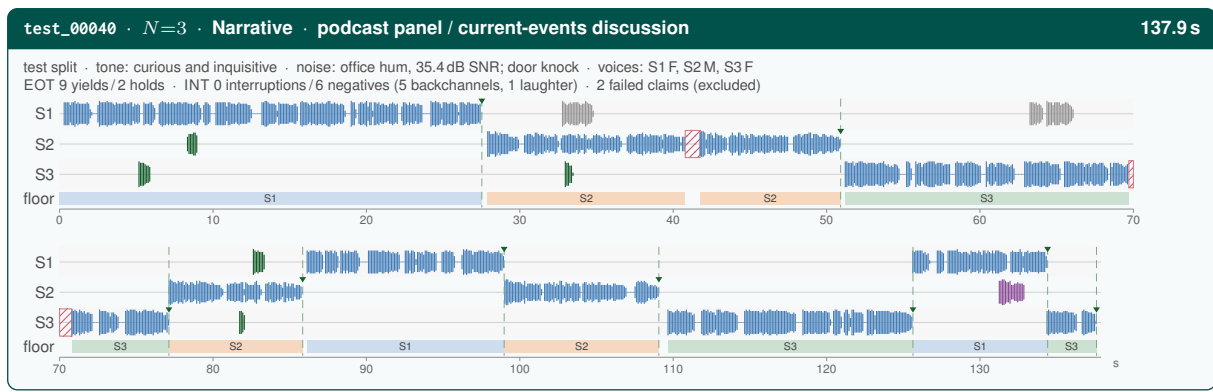
95.0–102.4 **S2** [turn] I wonder if that's just a phase, though. Like, do you think in five years we'll even distinguish between remote and office work? > yields to S1

102.4–113.4 **S1** [turn] That's tough to say. I suspect hybrid will stick around, just because it gives companies flexibility, but total remote—honestly, I think the novelty's worn off for a lot of people. > yields to S3

113.3–119.3 **S3** [turn] I'm not sure. I, uh, I still see arguments over Slack etiquette every week. Maybe some of it's just generational. > yields to S1

117.4–117.8 **S2** [backchannel] Yeah

119.3–126.2 **S1** [turn] Absolutely. Maybe next episode we can dive deeper into how leadership styles are adapting—there's a lot there. > yields, end



Transcript of test_00040

start—end (s) speaker [ground-truth label] text > outcome

0.3–15.6 **S1** [turn] All right, so I have to tell you both—yesterday I walked by the little park near my place and, no joke, there was a guy setting up his laptop on a tree stump and like, doing a video call with a parrot on his shoulder. Just out in the open, like it was totally normal.

5.2–5.9 **S3** [backchannel] That's wild.

8.3–9.0 **S2** [backchannel] No way!

15.7–27.5 **S1** [turn] And everyone passing by just kind of acted like okay, yeah, this is what public parks are for now, I guess—bird conference calls. I wanted to record it so badly but I was afraid he'd notice. > yields to S2

27.9–40.8 **S2** [turn] Honestly, that's so on-brand for this year. I was reading this article about how cities are like, kind of reshaping their public spaces because everyone started working outside? I mean, I even saw someone bring a monitor to the plaza downtown last week— > holds, pauses

32.8–34.8 **S1** [failed claim] Wait, like a whole monitor, just out there?

33.0–33.5 **S3** [backchannel] Oh wow—

41.8–50.9 **S2** [turn] —yeah, a full desktop monitor! They had an extension cord plugged into some weird little portable generator. It was the most committed outdoor office setup I've ever seen. > yields to S3

51.2–69.7 **S3** [turn] I mean, it raises all kinds of questions about access, right? Like, it's great people can work outside if they want, but that depends on having the tech, the power, even just space. I went to the botanical gardens last weekend and there were all these tiny pop-up tables suddenly, with people on calls. > holds, pauses

63.3–66.1 **S1** [failed claim] Yeah, exactly. And it, like, almost blurs the line—

70.8–77.1 **S3** [turn] —between public and private. That's what I keep thinking. Suddenly your office is everyone's background noise. > yields to S2

77.1–85.8 **S2** [turn] Right, and it's funny how people kind of stake out their own little territories, like you see someone with headphones and a coffee, and you just know, okay, that's their cubicle now. > yields to S1

81.7–82.1 **S3** [backchannel] Yeah.

82.6–83.4 **S1** [backchannel] Oh, for sure.

86.1–99.0 **S1** [turn] I tried it once, like, working in the park, but I couldn't focus at all. There was a dog zooming around chasing—what do you call it, those big bubble wands? It was amazing. But I ended up closing my laptop just to watch. > yields to S2

99.0–109.1 **S2** [turn] Haha, same here! I bring stuff to read and then, like, immediately get distracted by, I don't know, a marching band rehearsal or something totally random. Does that ever happen to you, Gina? > yields to S3

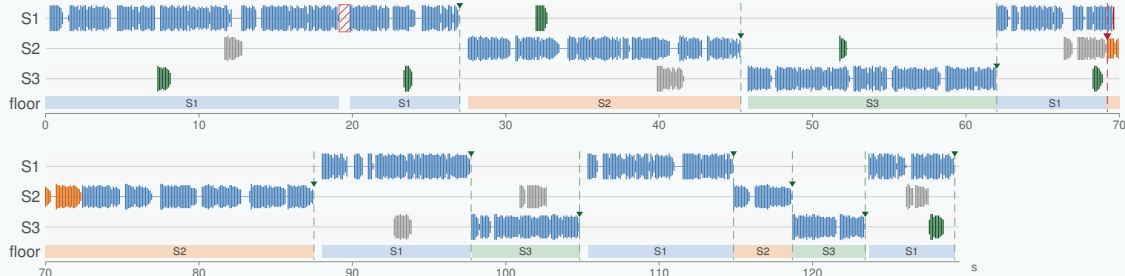
109.7–125.6 **S3** [turn] Honestly, I'm kind of the opposite. If I'm outside, I actually focus better. Maybe I tune out the chaos because I grew up in a noisy house. But I do wonder if we'll just adapt, you know—like, in the future, maybe it'll feel weird to work in total silence again. > yields to S1

125.6–134.4 **S1** [turn] That's such a good point. I mean—imagine office buildings piping in fake park sounds in five years just so we all feel like we're in the wild again! > yields to S3

131.3–133.0 **S2** [laughter] Haha, I'd actually love that.

134.4–137.6 **S3** [turn] Yeah, birdsong instead of hold music. I'm ready for that. > yields, end

test split · tone: professional and business-like · noise: cafe, 27.6 dB SNR · voices: S1 M, S2 F, S3 F
EOT 10 yields / 1 holds · INT 1 interruptions / 6 negatives (6 backchannels, 0 laughter) · 6 failed claims (excluded)



Transcript of test_00041

start-end (s) speaker [ground-truth label] text > outcome

0.3–19.1 S1 [turn] Morning, everyone. I thought I'd kick things off with a quick story from yesterday—I was walking through the customer support area and, uh, overheard two agents talking about how customers are still confused by the onboarding screens. I actually, um, sat in for a call and saw just how much folks struggle with that progress bar we added last sprint. > holds, pauses

7.3–8.2 S3 [backchannel] Oh really

11.7–12.9 S2 [failed claim] That echoes what I've heard—

19.9–27.0 S1 [turn] —like, it's supposed to give them reassurance but instead, it's making people anxious. I think we might need to rethink that entire flow. > yields to S2

23.4–23.9 S3 [backchannel] mm-hm

27.6–45.3 S2 [turn] So, I had a chat with Jackie from analytics last night, and she pulled up the numbers for drop-offs in onboarding. The weirdest part was that we see this spike in exits right at the, um, tutorial card with the animated checkmark. I, I thought people would love it, but apparently they just get stuck and quit. > yields to S3

32.0–32.8 S1 [backchannel] Interesting.

39.9–41.6 S3 [failed claim] That actually makes sense, yeah.

45.8–62.0 S3 [turn] I wanted to share something related—so last week, I did a sort of impromptu user test with my neighbor who's, um, not super tech-savvy. And she actually, like, couldn't get past that card. She thought she was supposed to click the animation, not the continue button. I ended up just sitting with her for fifteen minutes, helping her through every step. > yields to S1

51.8–52.3 S2 [backchannel] Oh wow.

62.0–69.6 S1 [turn] That's revealing—I mean, we probably overestimated how intuitive that would be. Maybe uh, we should be prototyping with even less guidance from us in the room. > cut off

66.4–69.1 S2 [failed claim] I agree. I—I think a stripped-down walkthrough could help.

68.3–68.9 S3 [backchannel] Right.

69.2–72.3 S2 [interruption] Actually—sorry, but I do want to push on that a little— > takes floor from S1

72.4–87.5 S2 [turn] I just worry if we take away too much structure, we'll get a different kind of confusion. Back when we tested with the older onboarding, people got lost because it was too open-ended. Maybe there's a way to do a softer introduction but not leave them unsupported. > yields to S1

88.0–97.7 S1 [turn] Yeah, you're right, the trick is balance. I guess, uh, we could run two prototypes in parallel, one with the progress bar and one with more descriptive text instead of the animation. > yields to S3

92.7–93.9 S3 [failed claim] mm-hm, I like that.

97.8–104.8 S3 [turn] And maybe—oh, um—maybe we can include an unobtrusive help button on that screen, so if they're lost, they have an option, but it doesn't get in the way? > yields to S1

100.9–102.7 S2 [failed claim] Yeah, that would cover both angles.

105.4–114.8 S1 [turn] Good idea. Alright, so next steps, I'll sync with Product to set up both flows for user testing. Does anyone want to help draft some short user prompts for the help button? > yields to S2

114.9–118.7 S2 [turn] I can take that on—just send me whatever context you want included. > yields to S3

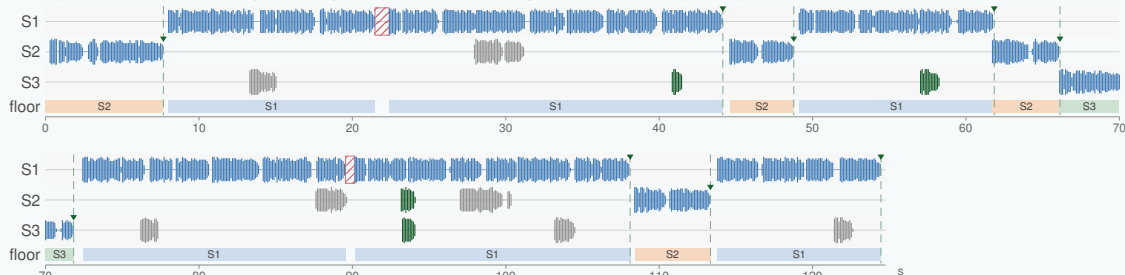
118.7–123.4 S3 [turn] Thanks, Sara. And thanks for flagging the customer story, Tom. That really brought the issues to life. > yields to S1

123.7–129.3 S1 [turn] Of course. Appreciate all the input, both of you. Let's check in next standup on early test results. > yields, end

126.1–127.6 S2 [failed claim] Great, looking forward to it.

127.6–128.6 S3 [backchannel] Sounds good.

test split · tone: professional and business-like · noise: outdoor park, 33.2 dB SNR; dishes clinking · voices: S1 M, S2 F, S3 M
EOT 9 yields / 2 holds · INT 0 interruptions / 4 negatives (4 backchannels, 0 laughter) · 7 failed claims (excluded)



Transcript of test_00049

start-end (s) speaker [ground-truth label] text > outcome

0.3–7.7 S2 [turn] All right, David, thanks for coming in. To start, can you walk us through a situation where you had to adapt quickly to a new challenge at work? > yields to S1

8.0–21.5 S1 [turn] Uh, yes, sure. So, um, last year when I was at LedgerTech, we suddenly lost two team members right before our product launch. I was still pretty new, and I—uh, I'd never led a project solo before. But, uh, I stepped up and took charge of the rollout plan for our new app feature. > holds, pauses

13.3–15.1 S3 [failed claim] Wow, right off the deep end, huh?

22.4–33.9 S1 [turn] Yeah, exactly. It was, uh, definitely a trial by fire. I had to reorganize the, um, deliverables, delegate what was left of the workload, and make sure we hit

the deadlines. I was terrified I'd mess something up, honestly.

28.0–31.2 **S2** [failed claim] How did you manage that anxiety, if you don't mind me asking—

34.0–44.2 **S1** [turn] —uh, right, so I kind of had to compartmentalize. I made lists, lots of lists, and reached out to a mentor for advice. It helped me stay focused on, like, the actual problems instead of just panicking. ▷ yields to S2

40.9–41.5 **S3** [backchannel] Smart move.

44.6–48.8 **S2** [turn] That sounds like a valuable lesson. So, did the launch go according to plan? ▷ yields to S1

49.1–61.8 **S1** [turn] Surprisingly, yeah. We, uh, actually managed to ship everything on time. There were some late nights, and a few small bugs we had to squash last minute, but, um, the manager praised the team for pulling it off. It was honestly a huge confidence boost for me. ▷ yields to S2

57.0–58.3 **S3** [backchannel] That's awesome, man.

61.7–66.1 **S2** [turn] Nice to hear. All right, let's switch gears. Mark, you wanted to ask about teamwork? ▷ yields to S3

66.1–71.8 **S3** [turn] Yeah! David, can you tell us about a time when you had to work with someone you didn't see eye to eye with? How'd you handle it? ▷ yields to S1

72.4–80.7 **S1** [turn] Oh. Uh, sure—I think the best example is from a group project at my previous job, where I had to work with someone who, uh, really liked to do things their own way.

76.2–77.4 **S3** [failed claim] Oh, I know the type.

80.9–89.5 **S1** [turn] Yeah, so, he would kind of dismiss my suggestions in meetings, and I, uh, I started to take it personally, which wasn't helping. So I asked him if we could talk one-on-one— ▷ holds, pauses

87.6–89.7 **S2** [failed claim] You initiated that conversation?

90.2–100.7 **S1** [turn] —yeah, I did. And it was super awkward at first, but we cleared the air, and I found out he actually respected my ideas, he just came off blunt without realizing. After that, things went a lot smoother.

93.2–94.1 **S2** [backchannel] Interesting.

93.3–94.1 **S3** [backchannel] That's a relief.

97.0–100.4 **S2** [failed claim] And did you two end up collaborating more after that, or—

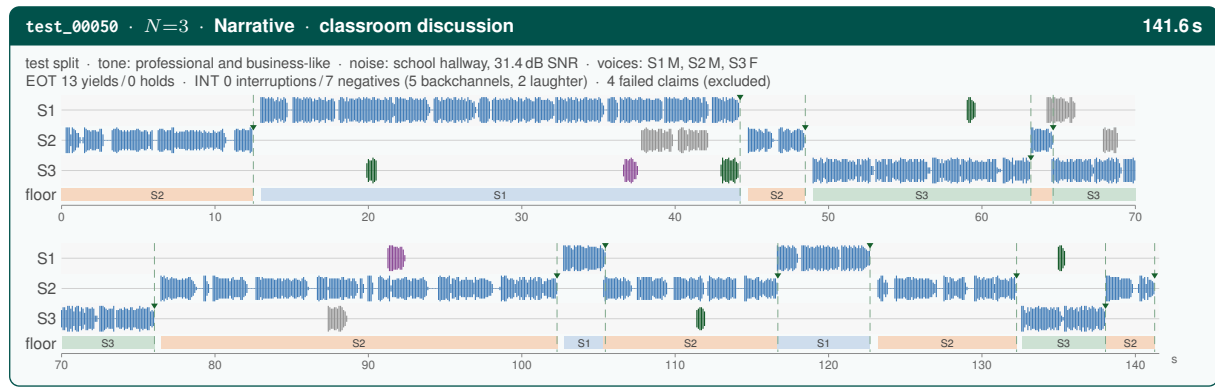
100.8–108.1 **S1** [turn] —yeah, actually, he became one of the people I trusted most on the team. He'd always give honest feedback, which I came to really appreciate. ▷ yields to S2

103.2–104.5 **S3** [failed claim] Love when that happens.

108.4–113.3 **S2** [turn] Great. Last question from me, David: Where do you see yourself growing if you joined us here at Verta? ▷ yields to S1

113.8–124.5 **S1** [turn] I, uh, I really want to deepen my technical skills, but I'm also excited about the chance to lead projects, maybe mentor junior teammates. I think this environment would push me in both those ways. ▷ yields, end

121.4–122.6 **S3** [failed claim] That's great to hear.



Transcript of test_00050

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–12.5 **S2** [turn] Alright everyone, thanks for joining today. We're focusing on memorable classroom moments—stories that stuck with you and maybe shaped the way you see teaching or learning. Who wants to start us off? ▷ yields to S1

13.0–32.1 **S1** [turn] Uh, I—I can go first, if that's okay. Um, so, one that comes to mind is from, I think, my sophomore year. I had this history teacher who would always, uh, come into class with, like, these weird props—one time he literally brought in a medieval helmet. And, uh, he passed it around while he talked about the Battle of Hastings. I just remember... I, uh, I actually paid attention that day, which was, you know, kind of rare for me in history.

19.9–20.6 **S3** [backchannel] Oh wow

32.2–39.3 **S1** [turn] Yeah, um, just holding that helmet, like, I guess it made it feel real? And he let us, uh, try it on, which was kinda embarrassing but also—

36.6–37.6 **S3** [laughter] haha right

37.8–42.2 **S2** [failed claim] Did he collect a bunch of stuff like that or was it just for that one lesson—

39.3–44.2 **S1** [turn] —no, he, uh, he actually had, like, a whole closet full. He joked it was his "history arsenal." ▷ yields to S2

43.0–44.1 **S3** [backchannel] That's amazing

44.8–48.5 **S2** [turn] That's a great example, thank you. Um, Sarah, do you have a story? ▷ yields to S3

49.0–63.2 **S3** [turn] Yeah, actually. So, um, mine is from elementary school—third grade, I think. My teacher, Ms. Bryant, she used to write personal notes on our assignments. And, uh, one time after I got a bad grade, she wrote, "I believe in you—don't give up." That really stuck with me, even now. ▷ yields to S2

59.0–59.6 **S1** [backchannel] mm-hm

63.2–64.6 **S2** [turn] That's—yeah, that's powerful. ▷ yields to S3

64.2–66.1 **S1** [failed claim] Did you, uh, keep the notes? Or—

64.6–69.1 **S3** [turn] —yeah, I actually still have a few of them in a scrapbook. It's kind of silly, I guess, but—

67.9–68.9 **S2** [failed claim] Not silly at all

69.3–76.1 **S3** [turn] —I've looked at them during some rough semesters, honestly. It reminded me that teachers notice things, even little ones. ▷ yields to S2

76.5–102.3 **S2** [turn] That actually reminds me of my own story. Um, mine's about a science class in high school. My teacher let us run the end-of-year project however we wanted. Two of us built this elaborate volcano model and, uh, it actually erupted in the middle of the hallway instead of the classroom. The principal walked by right as it was spewing fake lava everywhere. I thought we were in huge trouble, but she just laughed and took a picture. ▷ yields to S1

87.4–88.6 **S3** [failed claim] Oh I love that

91.3–92.4 **S1** [laughter] haha no way

102.7–105.5 **S1** [turn] Did it, um, get everywhere or just in one spot? ▷ yields to S2

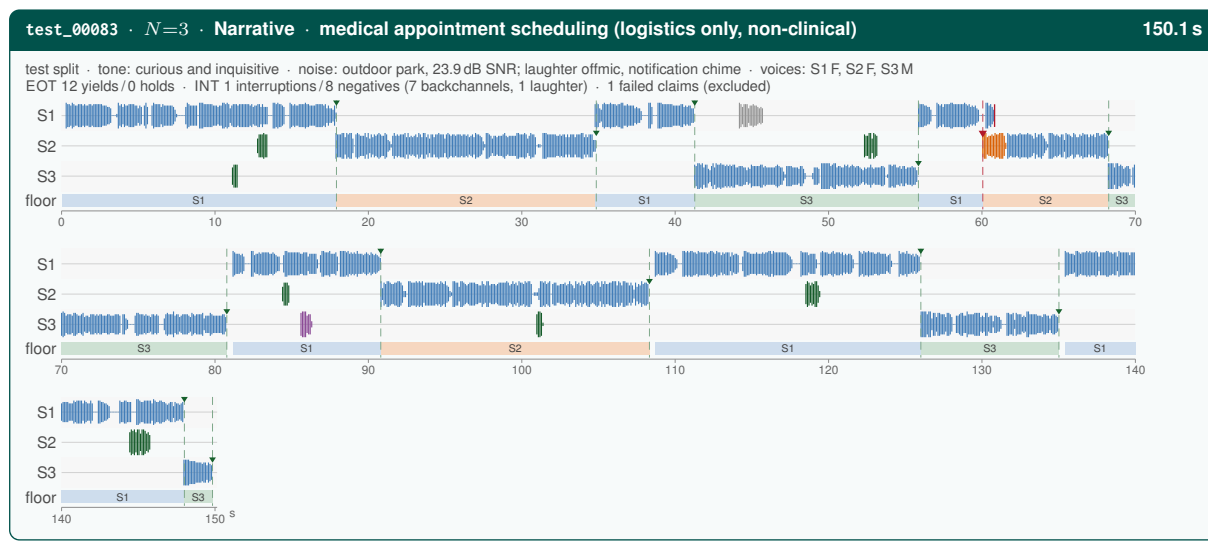
105.3–116.7 **S2** [turn] Mostly the tiles, thankfully. But we had to mop up together after. It kind of made me realize—like, the shared mess was the best part. Everyone was laughing, even the teacher. ▷ yields to S1

111.4–112.0 **S3** [backchannel] mm-hm

116.7–122.7 **S1** [turn] That's really, uh, cool. I guess, like, when teachers let things get a little messy, that's when the, um, fun stuff happens. ▷ yields to S2

123.2–132.3 **S2** [turn] Exactly. Sometimes the best lessons aren't the planned ones. Thanks both of you for sharing. Anyone have any last thoughts before we wrap up? ▷ yields to S3

132.6–138.1 **S3** [turn] Just—it's nice remembering these, honestly. Makes me want to keep a collection of good moments too. > yields to S2
135.0–135.5 **S1** [backchannel] yeah
138.1–141.3 **S2** [turn] Alright, that's a great note to end on. Thanks, everyone. > yields, end

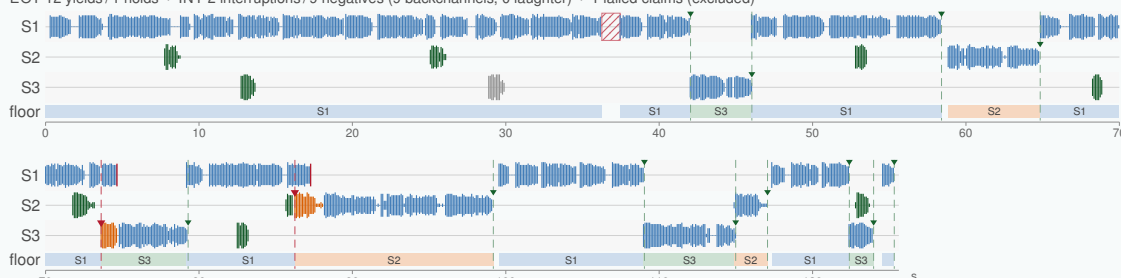


Transcript of test_00083

start–end (s) speaker [ground-truth label] text > outcome

0.3–17.9 **S1** [turn] You know, every time I try to schedule my annual checkup, it somehow turns into this, like, weird logistical puzzle. Last time, their online booking system wouldn't let me pick a date unless it was at least two months out, and even then, every slot was awkwardly in the middle of my work meetings. > yields to S2
11.1–11.5 **S3** [backchannel] mm-hm
12.8–13.5 **S2** [backchannel] oh wow
17.9–34.9 **S2** [turn] That's so frustrating. I actually had something similar last month. I called because the dentist's website showed nothing until July, but when I spoke to a real person, she just goes, 'Oh, we can squeeze you in next week if you don't mind a lunch hour.' It made me wonder if the online system is, I dunno, just a decoy. > yields to S1
34.8–41.3 **S1** [turn] Right, sometimes I wonder if they actually want us to call. Like, it's almost as if the tech is just for show. > yields to S3
41.3–49.5 **S3** [turn] I totally get that. So, I tried to book my physio appointment online and it gave me this error message that just said, like, 'Unknown system fault.' And then—
44.2–45.7 **S1** [failed claim] Oh, I hate those messages—
49.5–55.9 **S3** [turn] —so I ended up just walking into the clinic at nine in the morning, which is risky but, you know, sometimes you just need a human being. > yields to S1
52.3–53.2 **S2** [backchannel] exactly
55.9–60.8 **S1** [turn] Haha, brave. But it's true, you get way farther with an actual person. Once I— > cut off
60.0–61.6 **S2** [interruption] Did they see you right away? > takes floor from S1
61.7–68.3 **S2** [turn] Because honestly, I've thought about just dropping in, but I'm always worried they'll be, like, super annoyed or just send me away. > yields to S3
68.2–80.8 **S3** [turn] Yeah, that's what I figured too, but they actually just laughed and said people come in all the time because of those website bugs. They fit me in after a short wait. I think the receptionist appreciated having someone actually show up with a real face. > yields to S1
81.2–90.8 **S1** [turn] Makes sense. I feel like with medical scheduling, the more polite and apologetic you are, the better your odds, like you're negotiating access to some hidden club. > yields to S2
84.4–84.9 **S2** [backchannel] yeah
85.6–86.3 **S3** [laughter] haha true
90.8–108.3 **S2** [turn] It feels like a weird game. I remember once I had to get a referral for a specialist, but my doctor only did those on Tuesdays, and of course, the appointment desk closes early that day. So I ended up calling three different times, getting transferred each time, and finally some nurse overheard and just booked me manually. > yields to S1
101.0–101.4 **S3** [backchannel] wow
108.7–126.0 **S1** [turn] That sounds exhausting. I had something almost identical where I was bounced between departments, but the real trick is to say you have limited availability. I started saying, 'My schedule is really tight this month,' and suddenly they find hidden spots. It shouldn't have to be like that. > yields to S3
118.5–119.5 **S2** [backchannel] no kidding
126.0–135.0 **S3** [turn] That's a good tip! Next time I'm gonna try it. But do you think scheduling will ever actually get, like, easy? Or are we just destined to keep hacking our way through these systems forever? > yields to S1
135.4–148.0 **S1** [turn] Honestly, unless someone invents an app that can magically talk to all the clinics' old computers and charm the schedulers, I'm skeptical. But who knows, maybe we'll be able to book everything by text in a few years. > yields to S3
144.4–145.8 **S2** [backchannel] yeah, fingers crossed
148.0–149.8 **S3** [turn] Haha, I'll believe it when I see it. > yields, end

test split · tone: professional and business-like · noise: studio room tone, 27.3 dB SNR; notification chime · voices: S1 F, S2 M, S3 F
EOT 12 yields / 1 holds · INT 2 interruptions / 9 negatives (9 backchannels, 0 laughter) · 1 failed claims (excluded)



Transcript of test_00097

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–19.4 **S1** [turn] Thank you for joining the call. I just want to get everyone aligned—so the reason I reached out for tech support is that our CRM keeps timing out. It started, ah, yesterday afternoon about three, and every time I try to pull up a customer profile, it lags for a good thirty seconds before eventually just erroring out.

7.8–8.8 **S2** [backchannel] Right, right.

12.7–13.7 **S3** [backchannel] Mm, got it.

19.6–36.3 **S1** [turn] I called our internal help desk, but all they did was suggest clearing the browser cache, which I did—multiple times, in fact. But the problem persisted, which is why I escalated to you, since we depend on this data for, honestly, basic client communications. ▷ holds, pauses

25.1–26.2 **S2** [backchannel] Yeah, that's—

28.9–29.9 **S3** [failed claim] Can I check—

37.5–42.0 **S1** [turn] —it's just not sustainable, especially with end-of-quarter reporting deadlines. ▷ yields to S3

42.0–46.0 **S3** [turn] Did the issue start after any big update or was it totally out of the blue? ▷ yields to S1

46.0–58.4 **S1** [turn] No updates that I noticed. We got a server notification on Monday morning, just a banner saying there'd be maintenance on the backend at 2 AM, but this only cropped up more than twenty-four hours later. ▷ yields to S2

52.8–53.6 **S2** [backchannel] Interesting.

58.8–64.8 **S2** [turn] So, just to clarify, is everyone in your division seeing the same lag, or is it isolated to only your account? ▷ yields to S1

64.9–74.7 **S1** [turn] It's most of us, I think. I asked around on our group chat and at least four other folks said they've had basically identical issues, though oddly, one team member said he had no problems at all. ▷ cut off

68.3–68.9 **S3** [backchannel] Weird.

71.8–73.2 **S2** [backchannel] Yeah, that's strange.

73.6–74.7 **S3** [interruption] I just want to ask— ▷ takes floor from S1

74.8–79.3 **S3** [turn] Has anyone been able to try accessing the CRM outside of your corporate VPN? ▷ yields to S1

79.2–87.3 **S1** [turn] I did, actually. I was desperate enough last night to tether my laptop to my phone and try logging in, but the issue was exactly the same. It doesn't seem to be VPN-related. ▷ cut off

82.5–83.2 **S3** [backchannel] Okay.

85.7–86.2 **S2** [backchannel] Hm.

86.3–88.1 **S2** [interruption] Let me actually jump in— ▷ takes floor from S3

88.2–99.2 **S2** [turn] I've seen a very similar pattern after database re-indexing jobs, uh, sometimes there's a lag before the system actually re-balances itself. Has anyone mentioned any admin tasks recently getting scheduled? ▷ yields to S1

99.6–109.0 **S1** [turn] Not to me. The only thing, like I said, was that banner about general maintenance, and I didn't get any specifics. Is there a way to tell if that's the culprit? ▷ yields to S3

109.0–115.0 **S3** [turn] Yeah, I can pull the logs for your tenant and cross-reference with any scheduled scripts. It'll take a few minutes. ▷ yields to S2

114.9–117.1 **S2** [turn] Thanks, that should shed some light. ▷ yields to S1

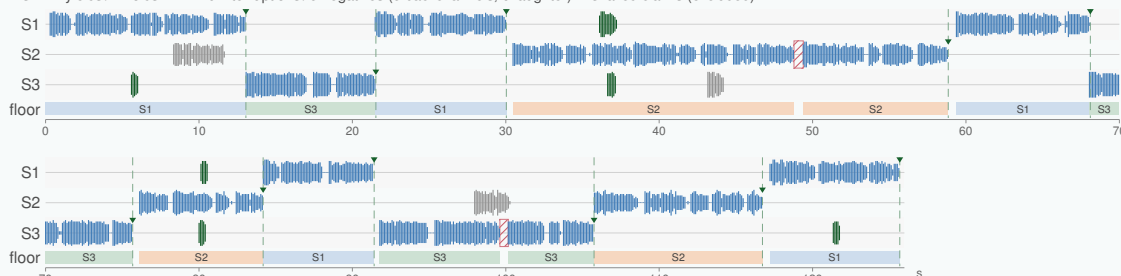
117.3–122.4 **S1** [turn] Appreciate it. I just hope it's something straightforward, because it's really holding everyone up. ▷ yields to S3

122.4–124.0 **S3** [turn] We'll get to the bottom of it. ▷ yields to S1

122.8–123.8 **S2** [backchannel] Absolutely.

124.6–125.3 **S1** [turn] Thank you both. ▷ yields, end

test split · tone: professional and business-like · noise: outdoor park, 35.2 dB SNR; laughter offmic, door knock · voices: S1 F, S2 M, S3 M
EOT 11 yields / 2 holds · INT 0 interruptions / 6 negatives (6 backchannels, 0 laughter) · 3 failed claims (excluded)



Transcript of test_00106

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–13.1 **S1** [turn] All right, let's get started. For *The Night Circus*, I'll just open by saying it actually surprised me—I thought the structure felt a bit fragmented, but I appreciated how Morgenstern played with timelines. Anyone else get thrown off by that at first? ▷ yields to S3

5.6–6.1 **S3** [backchannel] mm-hm

8.3–11.7 **S2** [failed claim] Yeah, I kept flipping back to the chapter headings to check the years.

13.1–21.5 **S3** [turn] I—well, actually, it took me a while to realize that the scenes weren't always chronological. Uh, I think it was maybe... halfway through, when Bailey's story started weaving in? ▷ yields to S1

21.5–30.0 **S1** [turn] Exactly, that's what I meant. I had to double-check myself too. But once I relaxed about it, it was kind of satisfying, watching it all click together as it went

along. ▷ yields to S2

30.5–38.2 **S2** [turn] Uh, it reminded me a bit of, you know, how Cloud Atlas is structured, with the nonlinear pieces. But I thought this one was less, uh, jarring somehow.

36.1–37.2 **S1** [backchannel] Yeah, agreed.

36.6–37.2 **S3** [backchannel] mm, right

38.4–48.8 **S2** [turn] But uh—I want to say the world-building here, the circus itself, is what stayed with me. I kept visualizing that ice garden. My daughter, she's fourteen, started sketching it after I described it at dinner. ▷ holds, pauses

43.2–44.3 **S3** [failed claim] Oh, that's great—

49.4–58.8 **S2** [turn] —yeah, and it really made me wish something like the circus could exist in real life, you know? It's evocative—the way she writes about smells, sounds... almost makes you forget there's a plot at all. ▷ yields to S1

59.4–68.1 **S1** [turn] That's interesting—because honestly I was less drawn to the circus atmosphere. I wanted more from Marco and Celia's relationship. I found their dynamic underdeveloped. ▷ yields to S3

68.0–75.7 **S3** [turn] I felt that too. There's something, uh, distant about their connection? It's like we're watching it behind glass. I wanted more dialogue or... I don't know, tension, maybe. ▷ yields to S2

76.1–84.2 **S2** [turn] Right, but I heard an interview with the author, and she said she wanted the romance to feel almost mythic, not realistic. Uh, she compared it to fairy tales. ▷ yields to S1

80.0–80.5 **S3** [backchannel] Right

80.1–80.6 **S1** [backchannel] mm-hm

84.2–91.4 **S1** [turn] Well, that makes sense. I suppose I'm just used to a bit more realism myself. Uh, can I ask—did anyone else think the ending was too abrupt? ▷ yields to S3

91.8–99.6 **S3** [turn] Actually, yes—I, uh, I was surprised when the last page came. I kept expecting another chapter, something to, um, tie up the Bailey thread more clearly. ▷ holds, pauses

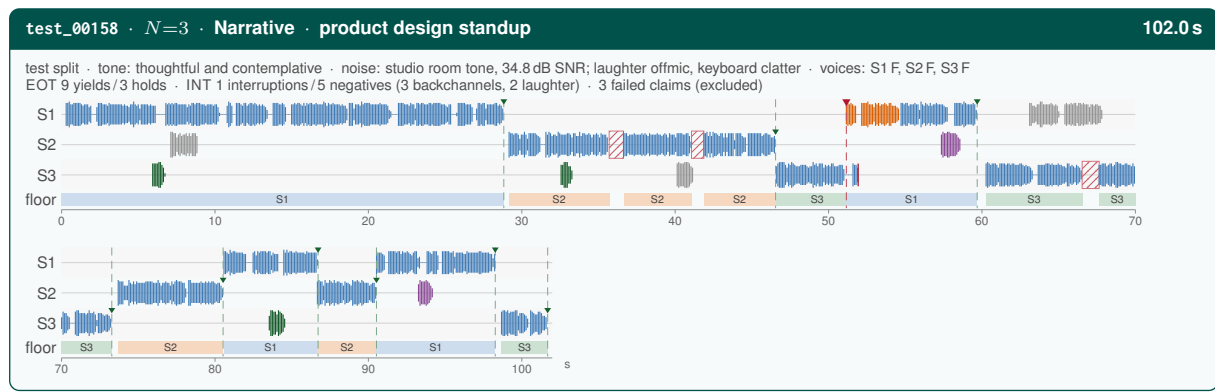
98.0–100.3 **S2** [failed claim] Well, I think the ambiguity was intentional—

100.2–105.8 **S3** [turn] —yeah, but even then, I just wanted a little more closure. Maybe a... a return to the present day, or something. ▷ yields to S2

105.8–116.7 **S2** [turn] I get it. Personally, I liked how it lingered in my mind after. Uh, that feeling of not quite leaving the circus behind. Reminded me of, you know—oh, what's that feeling when you finish a good show? Like you miss the world. ▷ yields to S1

117.2–125.7 **S1** [turn] Exactly. That's probably the best thing I can say for it—it made me want to reread parts, just to sort out the details, and that rarely happens for me with fiction. ▷ yields, end

121.4–121.8 **S3** [backchannel] Yeah



Transcript of test_00158

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–18.7 **S1** [turn] Alright, before we get into the bugs, can I just say yesterday's user testing was actually kind of hilarious? I had this one participant who tried to drag the new timeline widget—like, physically tried to swipe it off the screen, and when it didn't work, she gave me this look like I'd invented the flat tire of UI design.

5.9–6.8 **S3** [backchannel] Oh wow.

7.1–8.9 **S2** [failed claim] Yeah, it's confusing for sure.

18.8–28.8 **S1** [turn] Right? I just, I had to bite my tongue not to laugh. But honestly, it confirmed what we suspected about that affordance—people want to move it, but it's locked in place right now. ▷ yields to S2

29.2–35.7 **S2** [turn] That lines up with feedback from the client. They said their team tried to reorder milestones by dragging and assumed it was broken. ▷ holds, pauses

32.6–33.3 **S3** [backchannel] Uh-huh.

36.6–41.1 **S2** [turn] I mean, we either need to build that in now or at least give a signal you can't drag it yet. ▷ holds, pauses

40.1–41.2 **S3** [failed claim] Let me jump in—

41.9–46.5 **S2** [turn] —because right now it just kind of sits there like dead weight. We need an interim indicator at least. ▷ yields to S3

46.6–51.9 **S3** [turn] So, I watched three sessions from Tuesday, and every single person tried to reorder. Everyone. ▷ cut off

51.2–54.6 **S1** [interruption] Seriously? Even the guy with the, uh, the giant coffee mug? ▷ takes floor from S3

54.7–59.7 **S1** [turn] He looked like he was ready to fight the interface when nothing happened. It was like a silent stand-off. ▷ yields to S3

57.4–58.6 **S2** [laughter] Haha, I love that.

60.3–66.5 **S3** [turn] Yeah, him too. I think it's pretty cut-and-dried at this point. We need to prioritize this sprint for drag capability. ▷ holds, pauses

63.1–67.8 **S1** [failed claim] Wait, can we actually fit that in though? I know there's that blocker on the back-end ordering.

67.6–73.3 **S3** [turn] —We can if the API team delivers by Thursday. Otherwise, we'll have to fake it front-end for now. ▷ yields to S2

73.7–80.5 **S2** [turn] That works. Let's just make sure users get some kind of feedback, even if it's just a, I don't know, 'coming soon' tooltip or a shake animation. ▷ yields to S1

80.6–86.7 **S1** [turn] Oh, I love a good shake. Can I prototype that this afternoon? I have this idea with a little cartoon wiggle. ▷ yields to S2

83.5–84.6 **S3** [backchannel] Sounds good.

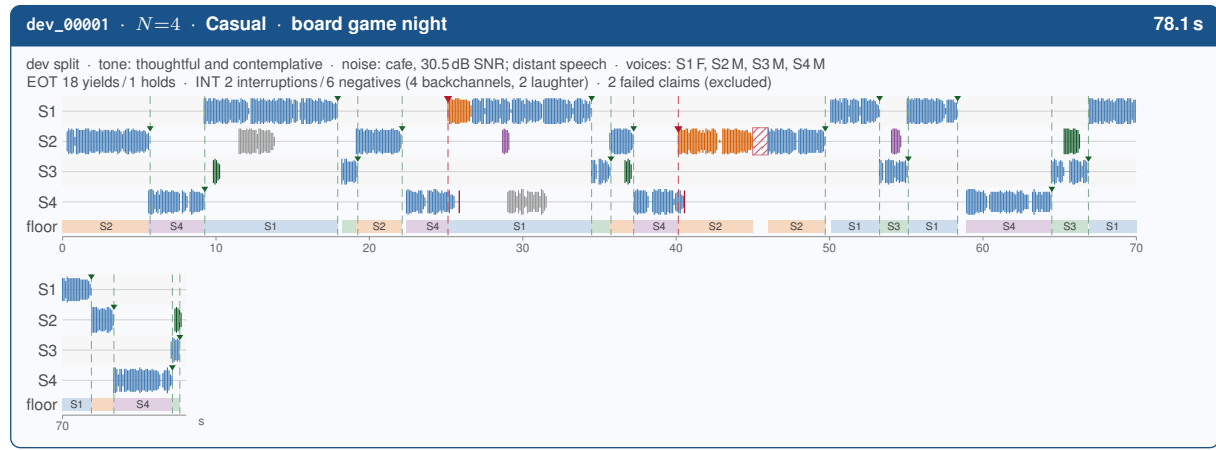
86.7–90.5 **S2** [turn] Perfect. If you run into anything, just ping me. I can rope in backend if needed. ▷ yields to S1

90.5–98.3 **S1** [turn] Awesome. I'll send a Slack when it's ready for feedback. And, uh, maybe warn the stakeholders their timelines might wiggle a little next week. ▷ yields to S3

93.3–94.2 **S2** [laughter] Haha, nice.

98.7–101.7 **S3** [turn] Alright, let's circle back tomorrow. Thanks, both of you. ▷ yields, end

K.3 Conversations with $N = 4$ speakers



Transcript of dev_00001

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–5.7 S2 [turn] So, uh, anyone remember the last time we actually finished a game of Settlers without arguing about the rules? ▷ yields to S4
5.6–9.3 S4 [turn] Never. Because every time you hoard the sheep, Dan, someone loses it. ▷ yields to S1
9.2–12.1 S1 [turn] That's actually a strategic move if you really think about it—
9.8–10.3 S3 [backchannel] mm-hm
11.5–13.8 S2 [failed claim] You mean if you're trying to make everyone miserable—
12.3–17.9 S1 [turn] —no, but listen, resource denial is valid, right? Like, isn't that the whole point of trading games? ▷ yields to S3
18.2–19.2 S3 [turn] Suppose it is. ▷ yields to S2
19.1–22.1 S2 [turn] I mean, that's technically true, but emotionally devastating. ▷ yields to S4
22.4–25.9 S4 [turn] Hey, speaking of trading, no one ever accepts my offers. I'm always tossing out sweet deals and just— ▷ cut off
25.1–26.6 S1 [interruption] Because you lowball everyone, Rick. ▷ takes floor from S4
26.7–31.3 S1 [turn] It's always like, 'Give me three wheat for a brick,' and then you act surprised we say no.
28.7–29.2 S2 [laughter] haha
29.0–31.6 S4 [failed claim] Three wheat for a brick is a completely fair—
31.4–34.5 S1 [turn] —no one in their right mind would take that. Even Matt wouldn't. ▷ yields to S3
34.5–35.8 S3 [turn] No, I would not. ▷ yields to S2
35.7–37.2 S2 [turn] See, democracy in action. ▷ yields to S4
36.7–37.1 S3 [backchannel] uh-huh
37.3–40.5 S4 [turn] Still, it's weird. No one ever targets Matt for the robber, even though he always wins. ▷ cut off
40.1–45.0 S2 [interruption] Honestly, that's because Matt just sits there all silent and then suddenly he's got like ten points. ▷ takes floor from S4 ▷ holds, pauses
46.0–49.7 S2 [turn] It's like stealth mode for board games. Nobody notices until it's too late. ▷ yields to S1
50.1–53.3 S1 [turn] That's actually kind of... impressive. I don't know how you do it. ▷ yields to S3
53.3–55.1 S3 [turn] Just boring enough to be ignored. ▷ yields to S1
54.0–54.7 S2 [laughter] haha, wow
55.1–58.3 S1 [turn] Right, but suddenly he's building cities all over. Like a ninja. ▷ yields to S4
58.9–64.5 S4 [turn] Yeah, okay, but this time—can we just agree, no secret port monopolies? Let's keep it transparent. ▷ yields to S3
64.5–66.9 S3 [turn] We can agree. Abiding is harder. ▷ yields to S1
65.3–66.4 S2 [backchannel] uh-huh, exactly
66.9–71.9 S1 [turn] Well if you want transparency, someone needs to actually read the rules out loud this time. ▷ yields to S2
71.9–73.3 S2 [turn] I volunteer as tribute. ▷ yields to S4
73.3–77.2 S4 [turn] Good. But if anyone starts the longest road debate again, I'm walking. ▷ yields to S3
77.1–77.6 S3 [turn] Noted. ▷ yields, end
77.3–77.8 S2 [backchannel] yeah



Transcript of dev_00042

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–7.9 S2 [turn] Alright, so, I figured we could start by just saying whether we actually finished the book this time—no pressure, but, uh, honesty hour? ▷ yields to S3

7.8–12.2 **S3** [turn] Oh, I did. Shockingly. Like, I was up late last night racing through those last chapters. > yields to S1

12.1–16.4 **S1** [turn] Wait, seriously? I'm... impressed, Mike. Last time you didn't make it past page forty. > yields to S4

16.5–19.9 **S4** [turn] I read the ending first. Skipped the middle as usual. > cut off

19.1–21.4 **S3** [interruption] No way, you just went straight to the last chapter? > takes floor from S4 > yields to S4

21.3–25.9 **S4** [turn] Yeah, like—I just want to know who survives. The rest is filler, honestly. > yields to S1

25.8–30.6 **S1** [turn] That's blasphemy, Pete. The build-up is the whole point. Otherwise none of the twists even land. > yields to S2

30.7–35.0 **S2** [turn] Exactly, thank you! The pacing is so intentional. He plants all those clues— > holds, pauses

31.9–32.4 **S3** [backchannel] Mhm

32.3–36.2 **S4** [failed claim] Or, you know, red herrings. The fish factory chapter, anyone?

32.6–33.0 **S1** [backchannel] Yeah

35.8–41.2 **S2** [turn] —and that's exactly why I loved the reveal at the fish factory. It felt set up and still got me. > yields to S1

39.8–41.4 **S3** [laughter] Haha, I totally fell for that.

41.5–47.6 **S1** [turn] I just—honestly, I was convinced the brother did it the whole time. Like, the little hints about his shoes, the late-night phone calls... > cut off

42.9–43.4 **S4** [backchannel] Nope.

44.2–44.9 **S2** [backchannel] Same here.

46.2–46.6 **S3** [backchannel] Uh-huh

46.7–50.9 **S3** [interruption] But wait, you guys didn't suspect the roommate? With all the weird locksmith references? > takes floor from S1

51.0–54.8 **S3** [turn] Like, why else would they keep dropping those hints about keys if it wasn't important? > yields to S1

54.8–58.5 **S1** [turn] But isn't that exactly what he wants you to think? It's almost too obvious. > yields to S2

58.4–66.0 **S2** [turn] Right, it's classic misdirection. But I will say, I did have to flip back a couple chapters to check the details about the safe. Like, was it actually locked the whole time or— > cut off

65.5–72.3 **S4** [interruption] No. There's a line where he says the safe is just decorative. Page ninety-eight. I remember because I was bored. > takes floor from S2

72.5–77.0 **S4** [turn] And I still don't believe it mattered. The author just likes red herrings. > yields to S2

76.9–79.6 **S2** [turn] Haha, you really have no faith in the process, Pete. > yields to S1

78.9–80.3 **S3** [laughter] Haha, classic Pete.

80.1–83.7 **S1** [turn] I mean, it keeps things interesting. Even if you jump to the end. > yields to S4

83.7–84.6 **S4** [turn] It's efficient. > yields to S3

85.0–89.9 **S3** [turn] So, next month—are we actually reading that sci-fi pick, or are we just picking another murder mystery? > yields to S2

90.0–92.1 **S2** [turn] Sci-fi. No skipping to the ending allowed. > yields to S1

92.1–93.7 **S1** [turn] We'll see. No promises. > yields to S3

93.6–95.9 **S3** [turn] Haha, I bet you'll all cave by chapter three. > yields, end



Transcript of test_00045

start—end (s) speaker [ground-truth label] text > outcome

0.3–10.5 **S2** [turn] Alright, so, um, welcome back everyone. I figured we could start with first impressions? Did anyone manage to finish the book or, like, at least most of it? > yields to S3

10.5–14.6 **S3** [turn] Yeah, I actually finished it last night, it was way more twisty than I expected! > yields to S1

14.6–23.9 **S1** [turn] Impressive. I made it to page one seventy-three before I got distracted by a documentary on snails. > yields to S4

23.8–29.7 **S4** [turn] Oh, uh, same here, well, not the snail part, but I didn't quite finish either. I'm about, I don't know, two-thirds through, maybe?

25.6–27.8 **S3** [failed claim] Did anyone else find all the footnotes kind of—

29.8–34.9 **S4** [turn] —distracting? Yes! I kept losing track of the main plot because I'd flip back and forth so much. > yields to S2

34.8–40.3 **S2** [turn] Right, right, I had to keep a finger wedged between pages just to find where I'd left off. > yields to S1

40.2–43.1 **S1** [turn] At least the footnotes weren't about snails. > yields to S2

40.9–41.4 **S4** [backchannel] mm-hm

42.0–42.5 **S3** [laughter] haha

43.4–49.9 **S2** [turn] So, uh, what did everyone think of the main character? Was anyone actually rooting for her? Because I, I kind of— > cut off

48.7–50.2 **S1** [interruption] —did not trust her at all. > takes floor from S2

50.2–58.3 **S1** [turn] Honestly, every chapter I was waiting for her to just betray someone. It felt inevitable. > yields to S4

58.3–64.1 **S4** [turn] Yeah, I mean, um, she kept hiding stuff from everyone—including the reader. It made me super anxious, honestly. > yields to S3

64.7–68.1 **S3** [turn] I thought it was kinda fun not knowing what she was up to, though. Like—

66.7–67.4 **S2** [backchannel] Right, right—

68.2–71.6 **S3** [turn] —I actually gasped at that scene with the key? You know, in the greenhouse? > yields to S2

71.5–75.1 **S2** [turn] Oh, absolutely. That whole reveal, I did not see it coming. > yields to S1

75.1–82.2 **S1** [turn] I was mostly impressed she found the key at all. I can't even find my socks in the morning. > yields to S2

77.2–77.8 **S4** [backchannel] right

79.5–80.4 **S3** [laughter] haha same

82.5–86.6 **S2** [turn] Did anyone think the pacing was kind of weird in the second half, though? Like, it really just— > cut off

85.5–87.1 **S3** [interruption] —slammed on the brakes, yeah! > takes floor from S2

87.2–91.2 **S3** [turn] There was all this set-up and then suddenly nothing happened for, like, fifty pages. > yields to S1

91.2–97.2 **S1** [turn] That's when the snail documentary started looking really good, not gonna lie. > yields to S2

96.0–96.6 **S4** [laughter] haha

97.5–100.4 **S2** [turn] I kinda liked the quiet parts, though? Is that weird? > yields to S4

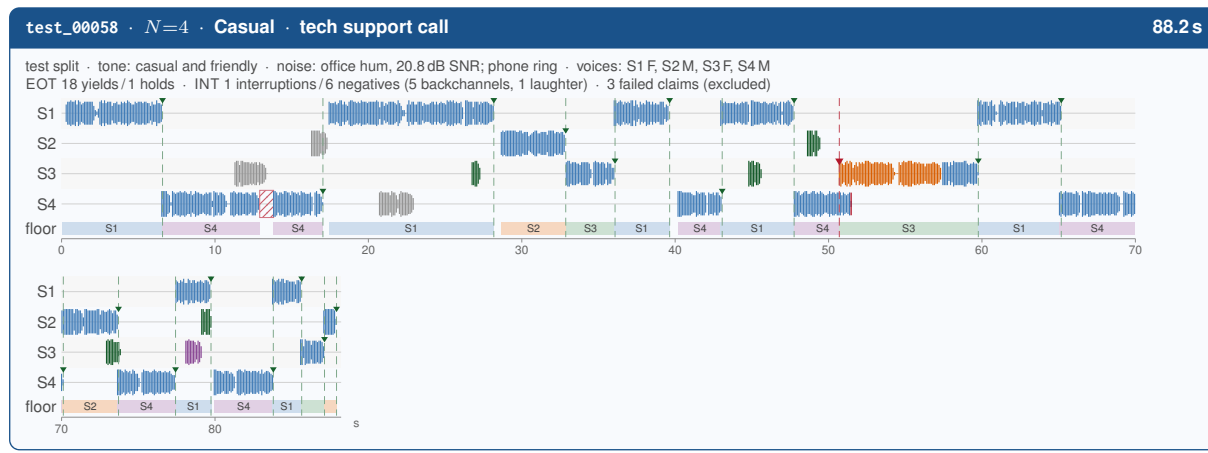
100.3–108.4 **S4** [turn] No, I get it. It gave me a break to, uh, catch my breath, right before that big twist at the end. Well, I haven't technically reached it yet, but people keep hinting. > yields to S3

108.3–109.8 **S3** [turn] Oh! Sorry, no spoilers! > yields to S1

109.8–114.1 **S1** [turn] We'll just sit here in ominous silence until he finishes. > yields to S2

111.4–111.8 **S3** [laughter] haha

114.6–116.2 **S2** [turn] Alright, so, next month's vote— > yields, end



Transcript of test_00058

start—end (s) speaker [ground-truth label] text > outcome

0.3–6.6 **S1** [turn] Hi everyone, thanks for hopping on. So, uh, Tom, you mentioned your Wi-Fi keeps dropping when you're trying to print, right? > yields to S4

6.5–12.9 **S4** [turn] Yeah, um, basically every time I try to print anything, it just says 'printer offline'. I, uh, I have no idea what's happening. > holds, pauses

11.3–13.4 **S3** [failed claim] Oh, that's so annoying, I hate when that happens—

13.8–17.0 **S4** [turn] —right? And I've tried restarting like, everything, but... nothing. > yields to S1

16.3–17.3 **S2** [failed claim] Yeah, that's rough man.

17.4–26.2 **S1** [turn] Did you happen to notice if the little Wi-Fi light on the printer is blinking or anything? Sometimes it just loses connection to the router for no reason.

20.7–23.0 **S4** [failed claim] Uh, I think it's solid? Hang on, let me—

26.3–28.2 **S1** [turn] —take a look, yeah. No worries. > yields to S2

26.8–27.3 **S3** [backchannel] Uh-huh.

28.7–32.9 **S2** [turn] Sometimes my printer just goes rogue and connects to the neighbor's Wi-Fi, I swear. > yields to S3

32.9–36.1 **S3** [turn] Haha, seriously, printers are like, the most rebellious gadgets. > yields to S1

36.0–39.6 **S1** [turn] For real! I had one that kept printing random receipts for weeks. > yields to S4

40.2–43.1 **S4** [turn] Okay, so the light is solid blue. That's good, right? > yields to S1

43.0–47.8 **S1** [turn] Yeah, that usually means it's connected, but uh, are you printing from your laptop or phone? > yields to S4

44.8–45.6 **S3** [backchannel] Good question.

47.8–51.5 **S4** [turn] Laptop. I, um, I tried from Chrome and from Word so I don't think it's the app? > cut off

48.6–49.5 **S2** [backchannel] Yeah, makes sense.

50.7–57.4 **S3** [interruption] Wait, do you have like, two printers set up in your settings by any chance? I totally did that once and it kept picking the wrong one. > takes floor from S4

57.5–59.8 **S3** [turn] Maybe it's just sending the print job into the void. > yields to S1

59.7–65.2 **S1** [turn] Oh, that's a good catch! Tom, can you check your printers list and see if there's like, two of the same model? > yields to S4

65.1–70.1 **S4** [turn] Uhh... wow, yeah, there are two. One says 'Copy 1' and the other's just the normal name. Uh, which do I...? > yields to S2

70.0–73.7 **S2** [turn] Try the one without 'Copy'? That's usually the, like, first installed one. > yields to S4

72.9–73.8 **S3** [backchannel] Yeah, exactly.

73.6–77.4 **S4** [turn] Okay, crossing my fingers. Annnd—oh! It's printing! Oh my god, yes! > yields to S1

77.5–79.7 **S1** [turn] Awesome, let's go! Glad we could help, Tom. > yields to S4

78.1–79.2 **S3** [laughter] Haha, team effort.

79.1–79.8 **S2** [backchannel] Nice, man!

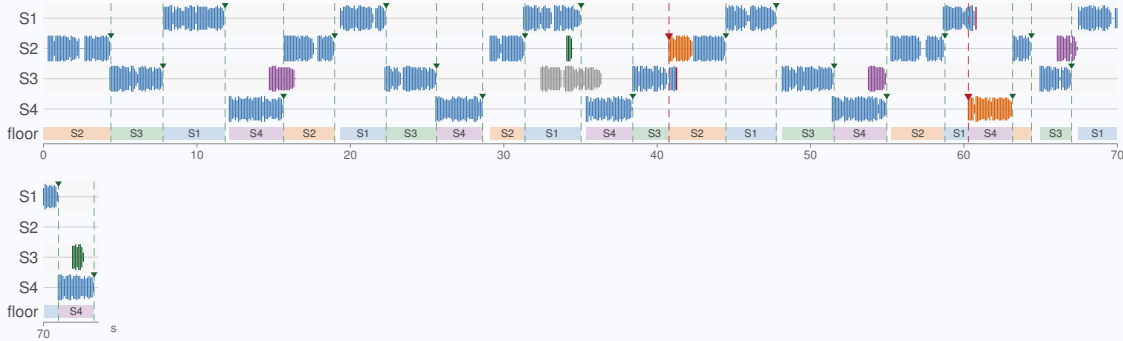
80.0–83.8 **S4** [turn] Thank you all, seriously. If it acts up again, can I, uh, bug you guys? > yields to S1

83.8–85.7 **S1** [turn] Anytime! That's what we're here for. > yields to S3

85.6–87.1 **S3** [turn] Yep, printer whisperers on call! > yields to S2

87.1–87.9 **S2** [turn] Have a good one, y'all. > yields, end

test split · tone: professional and business-like · noise: office hvac, 25.0 dB SNR; paper shuffle · voices: S1 M, S2 M, S3 F, S4 F
EOT 21 yields / 0 holds · INT 2 interruptions / 5 negatives (2 backchannels, 3 laughter) · 1 failed claims (excluded)

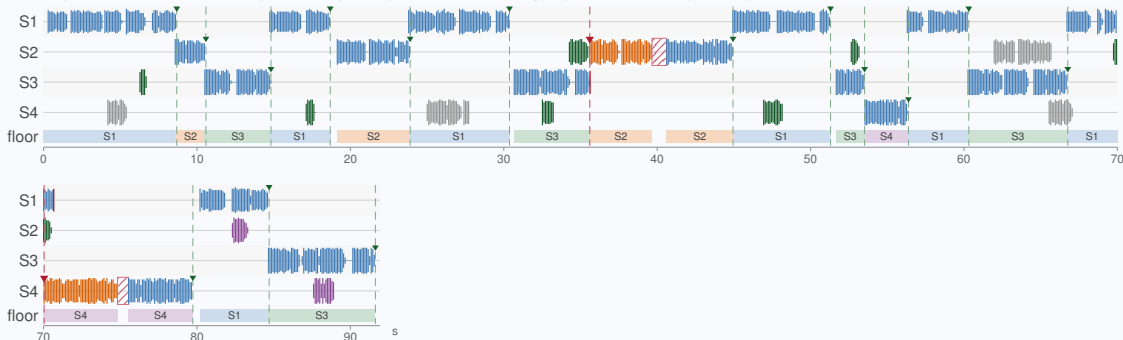


Transcript of test_00075

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–4.4 **S2** [turn] Alright, are we all set up for Codenames, or do we need to shuffle the deck again? ▷ yields to S3
 4.3–7.8 **S3** [turn] No, I already shuffled, like, twice! We're good there. Do you have the timer? ▷ yields to S1
 7.8–11.8 **S1** [turn] Uh, yeah, I think—it's... Let me check my phone, I, um, can put the timer on. ▷ yields to S4
 12.1–15.6 **S4** [turn] Let's just get started. We don't need a timer unless someone's super slow. ▷ yields to S2
 14.7–16.4 **S3** [laughter] Haha, that's calling us out already.
 15.7–19.0 **S2** [turn] No, I think a timer's smart, honestly. It keeps things moving. ▷ yields to S1
 19.4–22.3 **S1** [turn] Okay, uh, it's set for one minute per turn? Is that enough? ▷ yields to S3
 22.3–25.6 **S3** [turn] Yeah, that's perfect. And—oh, are we doing girls versus guys? ▷ yields to S4
 25.6–28.6 **S4** [turn] Or random teams? I mean, I don't want to crush you too fast. ▷ yields to S2
 29.1–31.4 **S2** [turn] Haha, ouch. Uh, I'm fine with anything really. ▷ yields to S1
 31.3–33.2 **S1** [turn] Yeah, um, maybe random? It might be—
 32.4–36.4 **S3** [failed claim] Wait, but if we do random, someone always complains they get the hard team.
 33.3–35.0 **S1** [turn] —uh, more fair though? But, it's up to you. ▷ yields to S4
 34.1–34.5 **S2** [backchannel] Mm-hm
 35.4–38.4 **S4** [turn] Let's just do guys versus girls, settle it like last time. ▷ yields to S3
 38.4–41.3 **S3** [turn] Right! Okay, who wants to be spymaster first? I like guessing more, so— ▷ cut off
 40.8–42.3 **S2** [interruption] I'll take spymaster, unless— ▷ takes floor from S3
 42.4–44.5 **S2** [turn] Unless you really want to do it, speaker_1? ▷ yields to S1
 44.5–47.8 **S1** [turn] Oh, no, I'm fine guessing. I always freeze up giving clues. ▷ yields to S3
 48.2–51.5 **S3** [turn] Same, honestly. Okay, speaker_4, are you spymaster for us? ▷ yields to S4
 51.4–55.0 **S4** [turn] Yeah, I can do spymaster. Just don't yell at me if I mess it up, please. ▷ yields to S2
 53.8–54.9 **S3** [laughter] Haha, no promises.
 55.2–58.8 **S2** [turn] Alright, so, speaker_1 and I are guessing. Ready for the first clue? ▷ yields to S1
 58.7–60.8 **S1** [turn] Uh, yeah—wait, who goes first? We do, right? Or— ▷ cut off
 60.3–63.2 **S4** [interruption] We flip for it, but you can just take the first turn if you want. ▷ takes floor from S1 ▷ yields to S2
 63.2–64.4 **S2** [turn] Thanks, that's efficient. ▷ yields to S3
 65.0–67.0 **S3** [turn] Alright, we're gonna destroy you. Just saying. ▷ yields to S1
 66.1–67.4 **S2** [laughter] Haha, we'll see about that.
 67.5–71.0 **S1** [turn] Okay, so, um, let's get started then. Who's got the word grid? ▷ yields to S4
 71.0–73.3 **S4** [turn] I've got it here. First clue coming in three... two... ▷ yields, end
 71.9–72.6 **S3** [backchannel] Ooh, suspense!

test split · tone: professional and business-like · noise: studio room tone, 22.1 dB SNR; chairs scraping; distant speech · voices: S1 F, S2 F, S3 M, S4 F
EOT 15 yields / 2 holds · INT 2 interruptions / 9 negatives (7 backchannels, 2 laughter) · 4 failed claims (excluded)



Transcript of test_00078

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.7 **S1** [turn] Okay, picture this—open shelving all along the kitchen wall, like something straight out of, I don't know, a magazine? Tons of plants, obviously. ▷ yields to S2
 4.2–5.5 **S4** [failed claim] Oh, I love that—
 6.3–6.7 **S3** [backchannel] Yeah
 8.6–10.6 **S2** [turn] And it keeps it so airy, honestly. ▷ yields to S3

10.5–14.8 **S3** [turn] But uh, would you actually keep it tidy? I mean, open shelving's pretty unforgiving. ▷ yields to S1

14.7–18.7 **S1** [turn] Rude! I mean, sure, it's chaos now, but if I get nice baskets, maybe? ▷ yields to S2

17.1–17.7 **S4** [backchannel] Mmm-hm

19.1–23.9 **S2** [turn] Oh totally, baskets are so helpful. And you could do matching jars for, like, pasta and stuff. ▷ yields to S1

23.8–25.5 **S1** [turn] Yes! Now you're speaking my language.

25.0–27.8 **S4** [failed claim] So are we thinking light wood for the shelves? Or—

25.6–30.4 **S1** [turn] —something darker maybe? I haven't really decided yet, honestly. What do you guys think? ▷ yields to S3

30.7–35.6 **S3** [turn] Lighter's gonna bounce way more light, especially if you keep the walls white. It'll feel bigger. ▷ cut off

32.5–33.3 **S4** [backchannel] Agreed.

34.3–35.5 **S2** [backchannel] Yes, exactly.

35.6–39.7 **S2** [interruption] Wait, can we talk counters for a second? I vote quartz, just for practicality. ▷ takes floor from S3 ▷ holds, pauses

40.6–44.9 **S2** [turn] I mean, it doesn't stain as much and honestly, I need something I can't ruin by accident. ▷ yields to S1

44.9–51.3 **S1** [turn] Haha, same, I cannot be trusted with marble. I dropped a beet on the test slab once, it was...not pretty. ▷ yields to S3

46.9–48.2 **S4** [backchannel] Oh no, really?

51.7–53.5 **S3** [turn] Yup, marble stains super fast. ▷ yields to S4

52.7–53.2 **S2** [backchannel] Right

53.6–56.4 **S4** [turn] Maybe a tiny butcher block section for, like, chopping? ▷ yields to S1

56.3–60.3 **S1** [turn] Oh, that's genius! But I want it built in, not just, like, plopped on top. ▷ yields to S3

60.3–62.5 **S3** [turn] Yeah, integrated is so much more seamless.

62.0–65.8 **S2** [failed claim] Wait, do we have space for that? I thought the layout was kinda tight by the stove—

62.7–66.8 **S3** [turn] —maybe by the window instead? There's that little bit, right where the light comes in. ▷ yields to S1

65.6–67.1 **S4** [failed claim] Mm, I love that spot.

66.7–70.7 **S1** [turn] Ooo, and I could put herbs there. Like, a mini herb garden right in the counter! ▷ cut off

69.7–70.5 **S2** [backchannel] Dreamy.

70.0–74.8 **S4** [interruption] Okay, so, moving on—do you want to keep the tile floor, or go for something warmer, like cork? ▷ takes floor from S1 ▷ holds, pauses

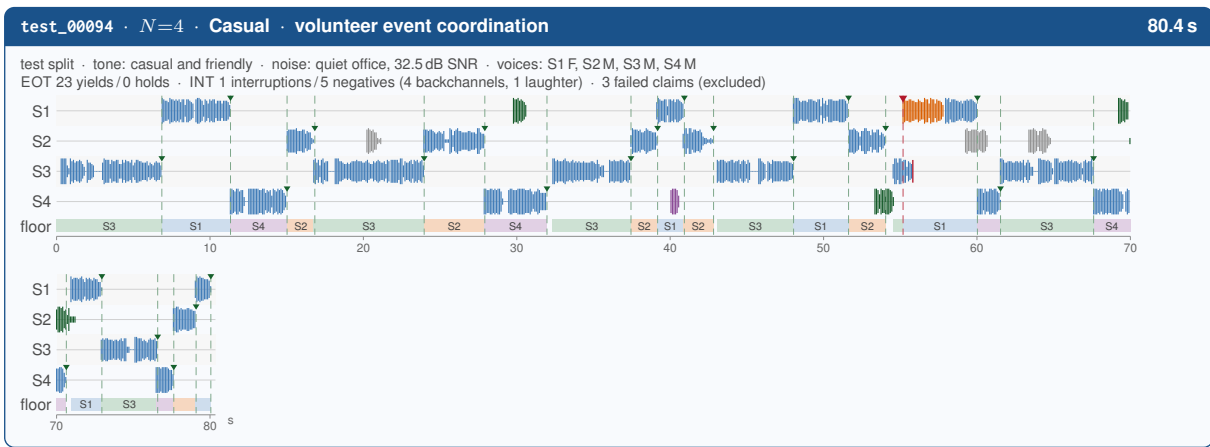
75.5–79.7 **S4** [turn] I just think cork feels so much softer underfoot if you're cooking for ages. ▷ yields to S1

80.2–84.7 **S1** [turn] Okay, you might be convincing me. My back is already grateful just thinking about it. ▷ yields to S3

82.3–83.3 **S2** [laughter] Haha, right?

84.7–91.6 **S3** [turn] Alright, so open shelves, light wood, quartz counters, cork flooring, and a window herb garden. Did we just design the whole thing? ▷ yields, end

87.6–89.0 **S4** [laughter] Haha, practically.



Transcript of test_00094

0.3–6.9 **S3** [turn] Alright, thanks for joining, everyone. Let's just figure out who's covering what for Saturday at the food pantry. ▷ yields to S1

6.9–11.3 **S1** [turn] I can do sign-in again, that's no problem. But I really can't stay past noon this time. ▷ yields to S4

11.3–15.0 **S4** [turn] Sounds good, and—I can handle packing boxes if that's still open? ▷ yields to S2

15.0–16.8 **S2** [turn] Yeah, uh, that'd help a lot, Paul. ▷ yields to S3

16.8–21.7 **S3** [turn] Perfect, thanks. Uh, Daniel, did you want to try doing the distribution line this time?

20.2–21.2 **S2** [failed claim] Well, so—

21.8–24.0 **S3** [turn] —I know it's a lot, but we've got folks to help. ▷ yields to S2

23.9–27.9 **S2** [turn] No, yeah, I can do it. Uh, I just might need a reminder on the procedure. ▷ yields to S4

27.9–32.0 **S4** [turn] Oh, it's not so bad. You get into a rhythm after the first few boxes. ▷ yields to S3

29.8–30.6 **S1** [backchannel] Exactly.

32.4–37.4 **S3** [turn] Just make sure everyone gets the same number of cans and veggies and, you know, smile a little. ▷ yields to S2

37.4–39.2 **S2** [turn] Haha, that'll be a stretch for me. ▷ yields to S1

39.1–40.9 **S1** [turn] Come on, you make grumpy look good. ▷ yields to S2

40.1–40.6 **S4** [laughter] Haha

40.8–42.8 **S2** [turn] Hey, thanks, I think? ▷ yields to S3

43.1–48.0 **S3** [turn] Okay, so, arrivals are at nine, right? Emily, can you bring the name tags again? ▷ yields to S1

48.0–51.6 **S1** [turn] I, uh—actually, can someone else do that? I keep forgetting mine in my car. ▷ yields to S2

51.7–54.0 **S2** [turn] I'll grab 'em. They're easy enough to keep track of. ▷ yields to S3

53.4–54.6 **S4** [backchannel] Right, thanks, Daniel.

54.5–55.8 **S3** [turn] Great. Uh, snack duty— ▷ cut off

55.2–57.8 **S1** [interruption] Actually, let me do snacks! I'll bake something this time. ▷ takes floor from S3

58.0–60.0 **S1** [turn] I felt bad bringing just fruit last time. ▷ yields to S4

59.3–60.7 **S2** [failed claim] Fruit was totally fine.

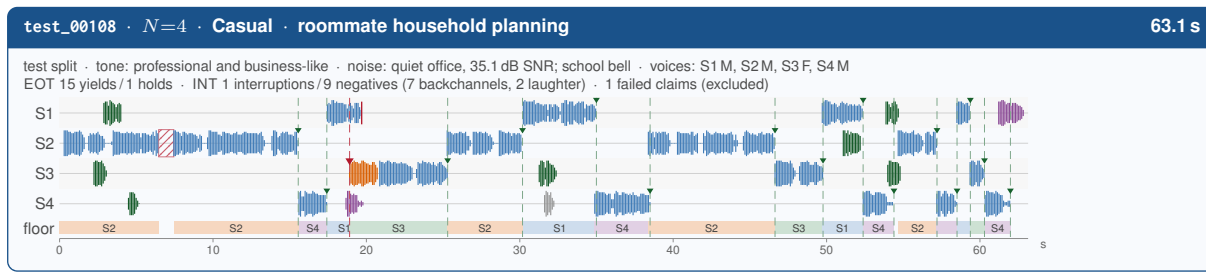
60.0–61.5 **S4** [turn] Yeah, but cookies sound better. ▷ yields to S3

61.5–65.0 **S3** [turn] Honestly, anything you make is an upgrade. Uh, last thing—

63.4–64.9 **S2** [failed claim] Oh, wait, one sec—

65.1–67.6 **S3** [turn] —Paul, do we need extra tables, or are we set? ▷ yields to S4

67.6–70.6 **S4** [turn] I'll bring two folding ones from the office, so we'll be fine. ▷ yields to S1
 69.2–69.9 **S1** [backchannel] Awesome.
 70.0–71.2 **S2** [backchannel] Yeah, perfect.
 70.9–73.0 **S1** [turn] Okay, anything else? Or can we call it? ▷ yields to S3
 72.9–76.6 **S3** [turn] Just text me if anything pops up. Otherwise, see you all at nine! ▷ yields to S4
 76.5–77.6 **S4** [turn] Looking forward to it. ▷ yields to S2
 77.6–79.1 **S2** [turn] Alright, see you all Saturday. ▷ yields to S1
 79.0–80.1 **S1** [turn] Bye everyone. ▷ yields, end



Transcript of test_00108

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.5 **S2** [turn] Alright, let's get started—weekly housekeeping. I know everyone's got stuff to do, so I'll keep this quick. ▷ holds, pauses
 2.2–3.1 **S3** [backchannel] Go ahead.
 2.9–4.1 **S1** [backchannel] Yeah, sounds good!
 4.5–5.2 **S4** [backchannel] Uh-huh.
 7.4–15.6 **S2** [turn] Okay, first—cleaning duties. So last week Alex did the kitchen, and Maria took the living room. Greg, you had the bathroom, right? ▷ yields to S4
 15.6–17.4 **S4** [turn] Yup. Got it done Saturday morning. ▷ yields to S1
 17.5–19.7 **S1** [turn] It was spotless, I noticed! Thanks, Greg. ▷ cut off
 18.7–19.8 **S4** [laughter] Heh, mm-hm.
 18.9–20.8 **S3** [interruption] Who's doing the kitchen this time? ▷ takes floor from S1
 20.9–25.3 **S3** [turn] I can swap and do it if no one minds. I'll trade if someone hates the vacuuming. ▷ yields to S2
 25.3–30.2 **S2** [turn] Actually, I was gonna offer! I kinda like vacuuming. Uh, so I'll take living room. ▷ yields to S1
 30.2–32.7 **S1** [turn] Yeah, I'll do the bathroom this week, that's fine.
 31.3–32.4 **S3** [backchannel] Great, thanks.
 31.6–32.3 **S4** [failed claim] Wait—
 32.7–35.0 **S1** [turn] —unless, Greg, you want the bathroom again? ▷ yields to S4
 34.9–38.5 **S4** [turn] Nah, you're up, Alex. I'll stick to trash detail this week, if that's cool. ▷ yields to S2
 38.4–46.6 **S2** [turn] Perfect, that's all set then. Um, next—supplies. We're definitely low on paper towels. Maria, didn't you grab them last time? ▷ yields to S3
 46.6–49.8 **S3** [turn] No, I did the dish soap. Paper towels was Alex. ▷ yields to S1
 49.7–52.4 **S1** [turn] Oh—right, that was me. Sorry! I can get more today. ▷ yields to S4
 51.1–52.2 **S2** [backchannel] Thanks, appreciate it.
 52.4–54.4 **S4** [turn] I'll pitch in for that, Alex. ▷ yields to S2
 53.9–54.7 **S1** [backchannel] Thanks, Greg.
 54.0–54.9 **S3** [backchannel] Same here.
 54.7–57.2 **S2** [turn] Alright—anything else for the week, or are we good? ▷ yields to S4
 57.2–58.5 **S4** [turn] Nope, that's it for me. ▷ yields to S1
 58.5–59.4 **S1** [turn] Yeah, I'm good! ▷ yields to S3
 59.4–60.3 **S3** [turn] All set. ▷ yields to S4
 60.3–62.0 **S4** [turn] Nice work, team. ▷ yields, end
 61.2–62.8 **S1** [laughter] Haha, so professional.

test split · tone: thoughtful and contemplative · noise: quiet apartment, 24.7 dB SNR; phone ring · voices: S1 M, S2 F, S3 M, S4 M
EOT 20 yields /3 holds · INT 3 interruptions /8 negatives (7 backchannels, 1 laughter) · 6 failed claims (excluded)



Transcript of test_00117

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–10.1 **S3** [turn] Alright so I figured, since we always end up talking about what we're watching or reading anyway, let's just make it official—book or movie recs, anyone? ▷ holds, pauses

7.7–9.7 **S2** [failed claim] Oh I love this, yes! Okay, can I—

11.0–12.4 **S3** [turn] —kick it off? Sure, jump in. ▷ yields to S2

12.4–21.5 **S2** [turn] Alright, so I just finished "Tomorrow, and Tomorrow, and Tomorrow." It totally got under my skin. Like, I kept thinking about all the choices the characters made, even after I put it down. ▷ cut off

20.6–23.4 **S4** [interruption] Oh, wait, I've heard of that one! Isn't it the gaming one? ▷ takes floor from S2

23.6–27.0 **S4** [turn] I keep seeing the cover everywhere, super colorful. Worth the hype? ▷ yields to S2

26.9–28.8 **S2** [turn] Yeah! Uh, honestly, I cried twice.

28.0–28.8 **S3** [backchannel] Wow, twice?

28.9–36.9 **S2** [turn] Twice! It's, um, really not just about games, though that's part of it. It's friendship and regret and how people change but don't, you know? ▷ yields to S1

37.5–41.7 **S1** [turn] That sounds...wow, that sounds pretty heavy. I, uh, I like those kinds of stories. ▷ yields to S3

38.9–40.2 **S4** [backchannel] Same here, honestly.

41.8–47.1 **S3** [turn] So you'd recommend it for, like, a chill weekend read or more an emotional investment thing? ▷ yields to S2

47.2–51.8 **S2** [turn] Definitely emotional investment. But it's still kinda fun? It's hard to describe. ▷ yields to S4

52.3–54.3 **S4** [turn] Oh, I'm adding it to my list, for sure. ▷ yields to S2

54.3–56.7 **S2** [turn] Haha yes! Okay, someone else go. Uh, Adam? ▷ yields to S1

57.2–63.2 **S1** [turn] Uh, okay. So I, um, I finally watched "The Last Black Man in San Francisco" last weekend. Have any of you seen it? ▷ holds, pauses

59.3–60.2 **S3** [failed claim] No, but I've—

64.1–74.3 **S1** [turn] —yeah, it's, I mean, it's not, like, plot-heavy? But the cinematography is gorgeous. It's kind of slow but really moving. Made me think about, uh, home and belonging. I keep replaying certain scenes in my head. ▷ holds, pauses

68.1–69.7 **S4** [failed claim] Same! Heard it's beautiful—

72.0–74.2 **S2** [failed claim] Oh, I've been meaning to watch that forever.

74.9–79.6 **S1** [turn] —yeah, it's quietly beautiful, I guess? Not flashy but really... honest. I dunno, it stuck with me. ▷ yields to S4

78.5–80.0 **S3** [failed claim] That's a strong endorsement.

79.9–82.5 **S4** [turn] Haha, he sold me. I'll add that to my list too. ▷ yields to S2

82.5–86.4 **S2** [turn] We're all gonna have, like, five new things on our lists by the end of this, aren't we? ▷ cut off

85.3–88.8 **S3** [interruption] That's basically why I wanted to do this. Alright, my turn— ▷ takes floor from S2

88.9–106.5 **S3** [turn] So, I read "Sea of Tranquility" by Emily St. John Mandel last week. Only took me two days. It's, um, quiet but resonant science fiction. Not at all what I expected—it weaves stories through time, but it's more about isolation and connection than, like, spaceships. ▷ yields to S4

94.2–94.6 **S2** [backchannel] Ooh

99.5–100.9 **S1** [failed claim] Oh I loved that, yeah.

106.9–108.7 **S4** [turn] I haven't even heard of it, actually. ▷ yields to S3

108.7–115.4 **S3** [turn] It sort of sneaks up on you. Super short, but you get attached to the characters anyway. The ending surprised me. ▷ yields to S4

115.9–119.3 **S4** [turn] Okay, putting it on the list—I'm gonna need a spreadsheet at this rate. ▷ yields to S1

117.3–118.1 **S2** [laughter] Haha, same.

119.2–120.7 **S1** [turn] So, uh, Josh, your pick? ▷ yields to S4

120.9–130.6 **S4** [turn] Yeah! Okay, I can't stop talking about "Everything Everywhere All at Once." It's just—like, it's so weird but so heartfelt? I laughed and then got weirdly existential. Has everyone seen it or...? ▷ yields to S2

125.3–126.2 **S3** [backchannel] Saw it, yeah.

126.6–127.4 **S1** [backchannel] Not yet.

130.6–132.8 **S2** [turn] I loved it! Bagel! The bagel, oh my god. ▷ yields to S4

132.8–134.8 **S4** [turn] Haha, yes! The bagel. Such a trip. ▷ cut off

133.9–135.9 **S1** [interruption] You're selling it, seriously. I need to— ▷ takes floor from S4

136.0–139.2 **S1** [turn] —finally get around to watching that one. Maybe this weekend. ▷ yields to S2

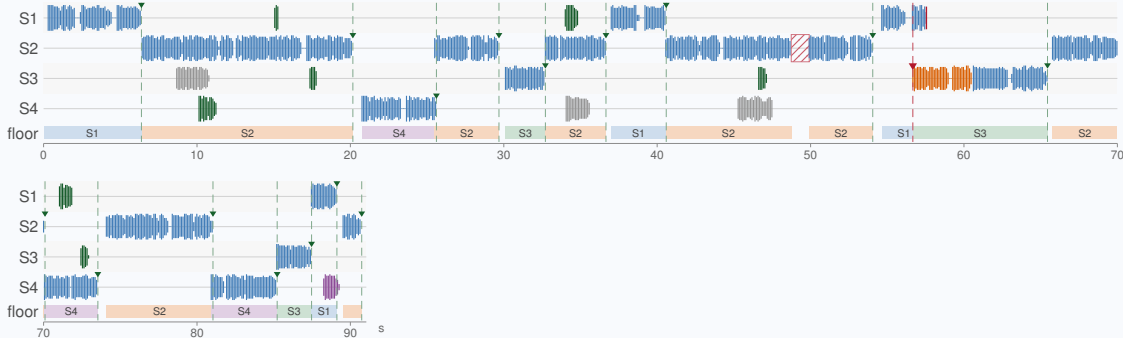
139.1–141.6 **S2** [turn] You have to! Then we can talk about googly eyes. ▷ yields to S3

141.7–146.9 **S3** [turn] Alright, I think we all just extended our to-watch and to-read lists by, like, a month. ▷ yields, end

145.2–146.0 **S4** [backchannel] Worth it, though.

145.6–146.6 **S2** [backchannel] Definitely worth it.

test split · tone: casual and friendly · noise: outdoor park, 37.8 dB SNR · voices: S1 M, S2 M, S3 M, S4 F
 EOT 16 yields / 1 holds · INT 1 interruptions / 8 negatives (7 backchannels, 1 laughter) · 3 failed claims (excluded)

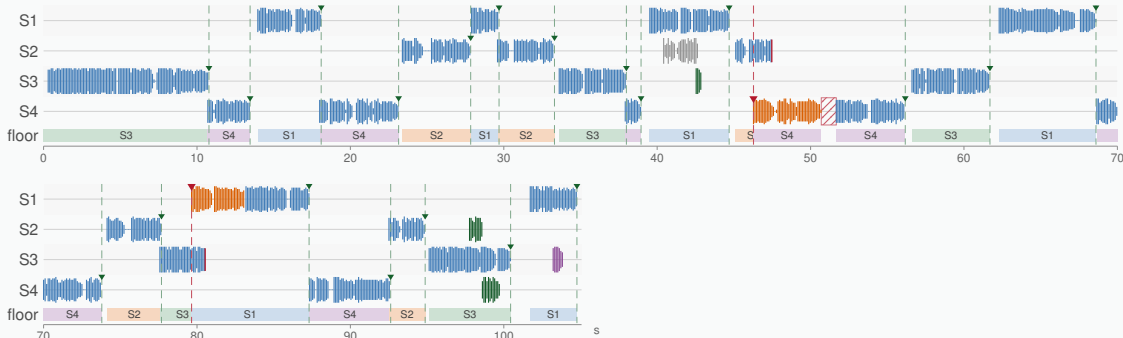


Transcript of test_00191

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.4 **S1** [turn] Alright, so, uh, Josh, you said you wanted some advice about, like, what happened with Clara last night? ▷ yields to S2
 6.4–12.4 **S2** [turn] Yeah, um, I mean, it's not a huge deal, just—I just felt weird, I guess. She kind of just... left suddenly?
 8.7–10.8 **S3** [failed claim] She just left, like, in the middle of dinner or—
 10.1–11.3 **S4** [backchannel] Oh no, really?
 12.6–20.2 **S2** [turn] —not quite, it was after, like, when we were just hanging out watching that dumb movie. Um, she said she was tired but it felt off, you know? ▷ yields to S4
 15.1–15.3 **S1** [backchannel] Mm.
 17.4–17.9 **S3** [backchannel] Yeah.
 20.8–25.6 **S4** [turn] I get that, sometimes when someone says they're tired it's not, it's not really about being tired, right? ▷ yields to S2
 25.5–29.7 **S2** [turn] Exactly. She seemed distracted all night. Like, her phone kept buzzing too. ▷ yields to S3
 30.1–32.7 **S3** [turn] Oh, was she texting someone? Or just, like, work stuff? ▷ yields to S2
 32.7–36.7 **S2** [turn] I—I don't know. Every time I glanced over she'd kind of hide her screen a bit? ▷ yields to S1
 34.0–34.9 **S1** [backchannel] Yeah, totally.
 34.1–35.6 **S4** [failed claim] That can feel really weird.
 37.0–40.6 **S1** [turn] So did you, like, ask her what was up? Or just kinda let it slide? ▷ yields to S2
 40.6–48.7 **S2** [turn] I didn't want to make it worse, so I just said, 'Hey, you alright?' and she said she was just tired from work, but then, like, immediately after, she left. ▷ holds, pauses
 45.3–47.6 **S4** [failed claim] Do you think maybe she's just stressed out? Or—
 46.6–47.1 **S3** [backchannel] Hmm.
 49.9–54.1 **S2** [turn] —yeah, it could be work, but I dunno, it felt... it felt personal somehow. ▷ yields to S1
 54.6–57.6 **S1** [turn] Like, do you think she's upset with you? Or is it just, like, her own stuff? ▷ cut off
 56.7–60.5 **S3** [interruption] Sometimes, um, people just need time, man. But also, if it's a pattern— ▷ takes floor from S1
 60.6–65.4 **S3** [turn] like, if this is something you notice a lot lately, then maybe you should actually ask more directly. ▷ yields to S2
 65.8–70.1 **S2** [turn] Yeah. I'm just afraid I'll push her away if I, like, come off as paranoid or needy. ▷ yields to S4
 70.1–73.5 **S4** [turn] But you deserve to know, too, you know? It's not necessary to care how she's feeling. ▷ yields to S2
 71.0–71.9 **S1** [backchannel] Yeah, for real.
 72.4–73.0 **S3** [backchannel] True.
 74.1–81.0 **S2** [turn] I think maybe I just—maybe next time it happens I'll just ask directly. Like, hey, is something up, not just about work. ▷ yields to S4
 80.9–85.2 **S4** [turn] That sounds good. And if you ever need to, like, run it by us before, you know where to find us. ▷ yields to S3
 85.2–87.5 **S3** [turn] Haha, yeah, group therapy Thursday nights. ▷ yields to S1
 87.5–89.1 **S1** [turn] We should start charging, honestly. ▷ yields to S2
 88.3–89.3 **S4** [laughter] Haha, right?
 89.5–90.7 **S2** [turn] Thanks, you guys, seriously. ▷ yields, end

dev split · tone: thoughtful and contemplative · noise: school hallway, 34.8 dB SNR; door knock, laughter offmic · voices: S1 F, S2 M, S3 F, S4 F
 EOT 20 yields / 1 holds · INT 2 interruptions / 4 negatives (3 backchannels, 1 laughter) · 1 failed claims (excluded)



Transcript of dev_00023

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–10.8 **S3** [turn] Alright, so if we're trying to pick a place for dinner this Friday, I think we should just, like, get the basic requirements out of the way. Dietary stuff, budget, location—let's list 'em. ▷ yields to S4
 10.7–13.5 **S4** [turn] Um, I don't eat red meat, just throwing that in first. ▷ yields to S1
 14.0–18.1 **S1** [turn] Is that like, no beef at all, or just like, you avoid it if there are other options? ▷ yields to S4

18.0–23.1 **S4** [turn] No beef, but I'll do chicken or fish, so it doesn't have to be full vegetarian or anything. ▷ yields to S2

23.4–27.8 **S2** [turn] Same for me, honestly. I can do most things but nothing too spicy. ▷ yields to S1

27.9–29.7 **S1** [turn] Wow, okay, so no Thai then? ▷ yields to S2

29.6–33.3 **S2** [turn] Yeah, I mean, I can do like one star heat but that's about it. ▷ yields to S3

33.6–38.0 **S3** [turn] Noted, no fire-breathing. What about allergies? Anything that's a non-starter? ▷ yields to S4

37.9–38.9 **S4** [turn] Nope, I'm good. ▷ yields to S1

39.5–41.4 **S1** [turn] I'm mildly lactose intolerant—

40.4–42.7 **S2** [failed claim] But pizza's still on the table, right?

41.6–44.7 **S1** [turn] —yeah, I just, uh, might need to ask for no cheese or something. ▷ yields to S2

42.5–42.9 **S3** [backchannel] mm-hm

45.1–47.5 **S2** [turn] So, like, is Italian the default then? ▷ cut off

46.3–50.7 **S4** [interruption] Or could we do Japanese? Sushi's not spicy and there's usually veggie options. ▷ takes floor from S2 ▷ holds, pauses

51.7–56.2 **S4** [turn] I know a place that does really good rolls, and they've got tempura and grilled stuff too. ▷ yields to S3

56.6–61.7 **S3** [turn] That's a good point. But is sushi gonna blow our budget? Last time we went it got kinda pricey. ▷ yields to S1

62.3–68.6 **S1** [turn] Yeah, I was gonna say, like, I'm fine with sushi but only if we're not ordering like, thirty things each. Maybe a set menu? ▷ yields to S4

68.7–73.8 **S4** [turn] Actually they have combos for like fifteen bucks if you want a couple rolls, it's not too bad. ▷ yields to S2

74.1–77.7 **S2** [turn] Sounds fine to me. I'll just avoid wasabi and I'm good. ▷ yields to S3

77.6–80.5 **S3** [turn] Okay, so Japanese is looking like the front runner unless there's an objection? ▷ cut off

79.6–83.1 **S1** [interruption] Wait, what about parking there? That lot is a nightmare on Fridays. ▷ takes floor from S3

83.2–87.3 **S1** [turn] If we take two cars, we won't find spots together. Should we Uber instead? ▷ yields to S4

87.3–92.6 **S4** [turn] Oh! I can drive, and I've got space for three if someone doesn't mind being a little squished in the back. ▷ yields to S2

92.5–94.9 **S2** [turn] That works. I'll call shotgun now. ▷ yields to S3

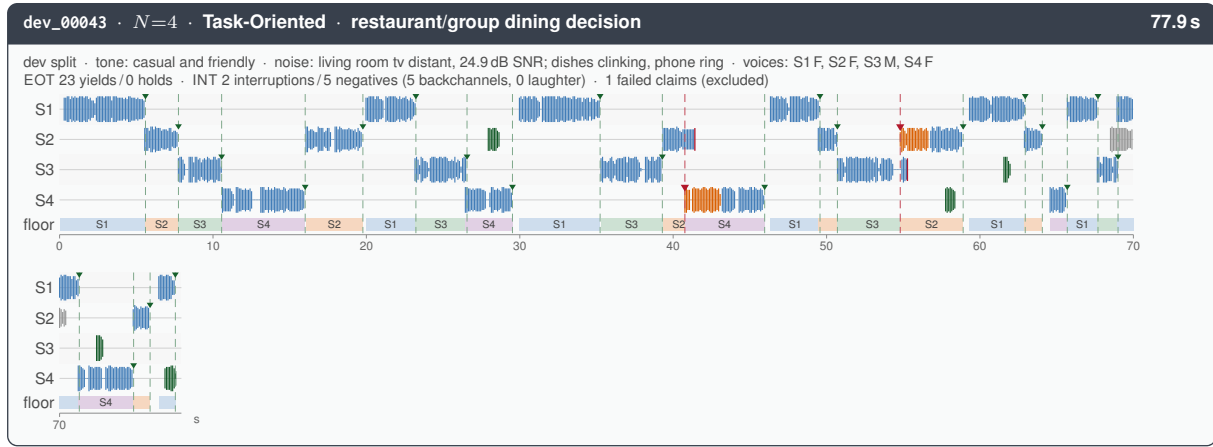
95.1–100.4 **S3** [turn] Alright, looks like we have a plan—Japanese, ride with Ava, Friday at seven. Everyone good? ▷ yields to S1

97.8–98.7 **S2** [backchannel] All good here.

98.6–99.8 **S4** [backchannel] Yep, sounds fun!

101.7–104.8 **S1** [turn] Yeah, sure. Let's see if we actually stick to this plan for once. ▷ yields, end

103.2–103.9 **S3** [laughter] haha right



Transcript of dev_00043

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–5.6 **S1** [turn] Alright, so let's start with options—do we want to go for Italian again or try something new tonight? ▷ yields to S2

5.5–7.7 **S2** [turn] Oh, I kind of want sushi honestly. ▷ yields to S3

7.7–10.6 **S3** [turn] Same. I had pasta three times last week. ▷ yields to S4

10.6–16.0 **S4** [turn] If we're voting, I'm good with sushi. But let's check if everyone can actually get something they'll eat— ▷ yields to S2

16.1–19.8 **S2** [turn] Yes, please! I don't want to, like, accidentally veto anyone's dinner. ▷ yields to S1

20.0–23.2 **S1** [turn] I'm fine with sushi, as long as there's something vegetarian. ▷ yields to S3

23.2–26.6 **S3** [turn] Most places have avocado rolls or something, right? ▷ yields to S4

26.5–29.5 **S4** [turn] Yeah, and cucumber rolls. Worst case, edamame. ▷ yields to S1

27.9–28.7 **S2** [backchannel] Oh, yeah.

30.0–35.2 **S1** [turn] Okay, that's settled. Does anyone know a spot nearby that's not totally packed on Fridays? ▷ yields to S3

35.3–39.3 **S3** [turn] Miso Place usually has tables open, but it's kind of slow. ▷ yields to S2

39.3–41.4 **S2** [turn] Oh but the teriyaki is really good there. ▷ cut off

40.8–43.1 **S4** [interruption] Wait, is that the one with the lanterns all over? ▷ takes floor from S2

43.2–46.0 **S4** [turn] Because if it is, they have a decent veggie menu too. ▷ yields to S1

46.3–49.6 **S1** [turn] Yeah, I think that's it. Anyone actually checked their website? ▷ yields to S2

49.5–50.7 **S2** [turn] Hang on, let me look. ▷ yields to S3

50.7–55.3 **S3** [turn] I mean, I'm good with anywhere as long as they don't close early. Last time— ▷ cut off

54.8–56.7 **S2** [interruption] Wait, they're open 'til ten tonight! ▷ takes floor from S3

56.8–58.9 **S2** [turn] So, yeah, we should be totally fine. ▷ yields to S1

57.8–58.4 **S4** [backchannel] Perfect.

59.3–63.0 **S1** [turn] Alright, let's aim for eight? That gives everyone time to get over there. ▷ yields to S2

61.6–62.0 **S3** [backchannel] Mm-hm.

62.9–64.1 **S2** [turn] Sounds good to me. ▷ yields to S4

64.6–65.7 **S4** [turn] Is anyone driving? ▷ yields to S1

65.7–67.7 **S1** [turn] I can, but I only have space for two more. ▷ yields to S3

67.6–69.0 **S3** [turn] I'll probably just bike. ▷ yields to S1

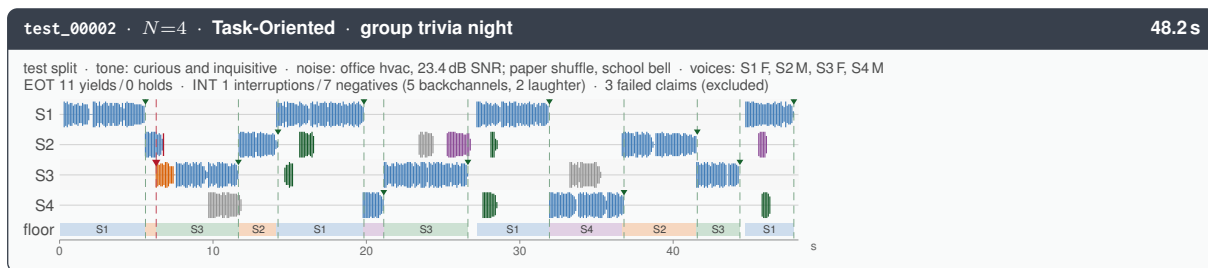
68.5–70.5 **S2** [failed claim] Can I grab a ride with you, Alex?

68.9–71.3 **S1** [turn] —Yeah, no problem. I'll swing by your place first. ▷ yields to S4

71.2–74.8 **S4** [turn] Awesome, I'll meet you all there and grab a table if I get in early. ▷ yields to S2

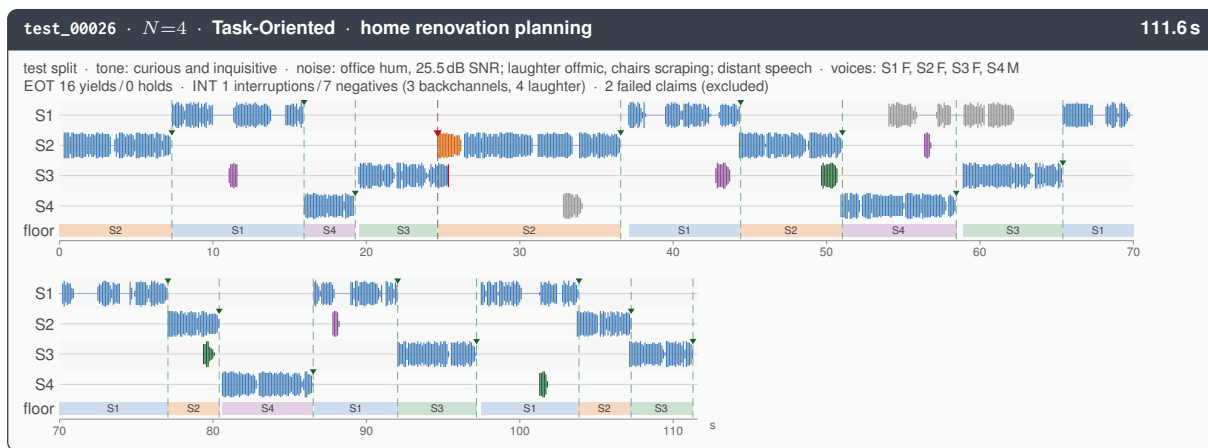
72.4–72.9 **S3** [backchannel] Deal.

74.8–75.9 **S2** [turn] Yay, sushi night. ▷ yields to S1
 76.5–77.5 **S1** [turn] Alright, see you all at eight. ▷ yields, end
 76.9–77.6 **S4** [backchannel] See you then.



Transcript of test_00002 start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–5.6 **S1** [turn] Alright, round three just started. The first question's about world capitals—what's the capital of New Zealand? ▷ yields to S2
 5.6–6.8 **S2** [turn] Oh, that's easy, it's Auckland— ▷ cut off
 6.3–7.5 **S3** [interruption] Actually, wait— ▷ takes floor from S2
 7.6–11.7 **S3** [turn] I think it's Wellington, not Auckland. Auckland's just the biggest city. ▷ yields to S2
 9.7–11.9 **S4** [failed claim] Yeah, Wellington. That's definitely right.
 11.7–14.2 **S2** [turn] Oh man, you're totally right. I always mix those up. ▷ yields to S1
 14.2–16.3 **S1** [turn] Okay, so final answer: Wellington?
 14.7–15.3 **S3** [backchannel] Yeah.
 15.7–16.7 **S2** [backchannel] Wellington for sure.
 16.4–19.8 **S1** [turn] Great. Next one says, "What element has the chemical symbol 'Sn'?" ▷ yields to S4
 19.8–21.1 **S4** [turn] That's, uh... tin, right? ▷ yields to S3
 21.2–24.0 **S3** [turn] Yeah, it's tin. Sn is from the Latin name, stannum.
 23.4–24.4 **S2** [failed claim] Smart! Never would've—
 24.1–26.6 **S3** [turn] —guess that if I hadn't taken chem last year. ▷ yields to S1
 25.3–26.8 **S2** [laughter] Yeah, real chem genius over here.
 27.2–28.8 **S1** [turn] Alright, we locking in tin?
 27.6–28.5 **S4** [backchannel] Absolutely.
 28.1–28.6 **S2** [backchannel] Yep.
 28.8–31.9 **S1** [turn] Okay, third question: Who painted The Persistence of Memory? ▷ yields to S4
 32.0–33.7 **S4** [turn] Dali. That's Salvador Dali—
 33.3–35.3 **S3** [failed claim] Wait, is it Dali or Magritte?
 33.8–36.8 **S4** [turn] —no, I'm pretty sure that one's Dali. The melting clocks? ▷ yields to S2
 36.7–41.6 **S2** [turn] Yeah, melting clocks, that's Dali. I remember that poster everywhere in college dorms. ▷ yields to S3
 41.5–44.3 **S3** [turn] Right, okay, that's definitely Dali. Good call. ▷ yields to S1
 44.7–47.9 **S1** [turn] Alright, I'll write down Dali for that one. Ready for the next question? ▷ yields, end
 45.6–46.1 **S2** [laughter] Bring it on.
 45.8–46.4 **S4** [backchannel] Ready.



Transcript of test_00026 start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–7.3 **S2** [turn] Alright, so should we start with the kitchen or the living room? I feel like the kitchen is probably the bigger headache, but I'm up for either. ▷ yields to S1
 7.4–15.9 **S1** [turn] Let's rip off the kitchen band-aid first. If I have to look at those cabinets any longer, I might stage a coup. ▷ yields to S4
 11.1–11.7 **S3** [laughter] haha
 16.0–19.3 **S4** [turn] Seriously though, what's your dream look, like, minimal or more farmhouse chic? ▷ yields to S3
 19.5–25.3 **S3** [turn] So, um, I'm actually really curious—what's our budget ceiling for the kitchen? Just so we don't go too wild with the Pinterest boards. ▷ cut off
 24.6–26.2 **S2** [interruption] Oh! Wait, I did the math last night— ▷ takes floor from S3
 26.4–36.6 **S2** [turn] Honestly, including appliances, we should probably not go over fifteen grand. I mean, unless we suddenly win the lottery. But I found some pretty solid options online. ▷ yields to S1
 32.9–34.1 **S4** [failed claim] Yeah, yeah, that makes sense.
 37.1–44.4 **S1** [turn] Fifteen grand. So just enough for one marble countertop and a sense of regret. ▷ yields to S2

42.8–43.7 **S3** [laughter] haha right

44.3–51.0 **S2** [turn] What if we do butcher block, though? It looks good and it's, like, a quarter of the price. Maybe then we'd have money for the backsplash. $\triangleright$ yields to S4

49.7–50.7 **S3** [backchannel] Yeah, totally.

50.9–55.1 **S4** [turn] I do like the look of that. And it's not, you know, as cold as marble in the winter.

54.1–62.3 **S1** [failed claim] But will you actually maintain it? No offense, but I saw what happened to the cutting board last year—

55.2–58.5 **S4** [turn] —Alright, that's fair. I promise I'll learn from my mistakes. Maybe. $\triangleright$ yields to S3

56.4–56.8 **S2** [laughter] haha

58.9–65.4 **S3** [turn] Okay, so, for layout—are we changing the footprint at all, or just swapping finishes? Because that's a huge cost jump. $\triangleright$ yields to S1

65.4–77.1 **S1** [turn] If we keep the plumbing where it is, we can funnel more money into lighting. My one demand is that we aren't cooking in the dark anymore. $\triangleright$ yields to S2

77.1–80.4 **S2** [turn] Oh yeah, those fluorescent tubes are, like, a crime against humanity. $\triangleright$ yields to S4

79.4–80.1 **S3** [backchannel] Yup.

80.6–86.5 **S4** [turn] I can install new pendants, honestly. I watched a video and I think I won't electrocute myself. Probably. $\triangleright$ yields to S1

86.6–92.0 **S1** [turn] Great, so step one: put life insurance paperwork on the fridge. $\triangleright$ yields to S3

87.8–88.2 **S2** [laughter] haha

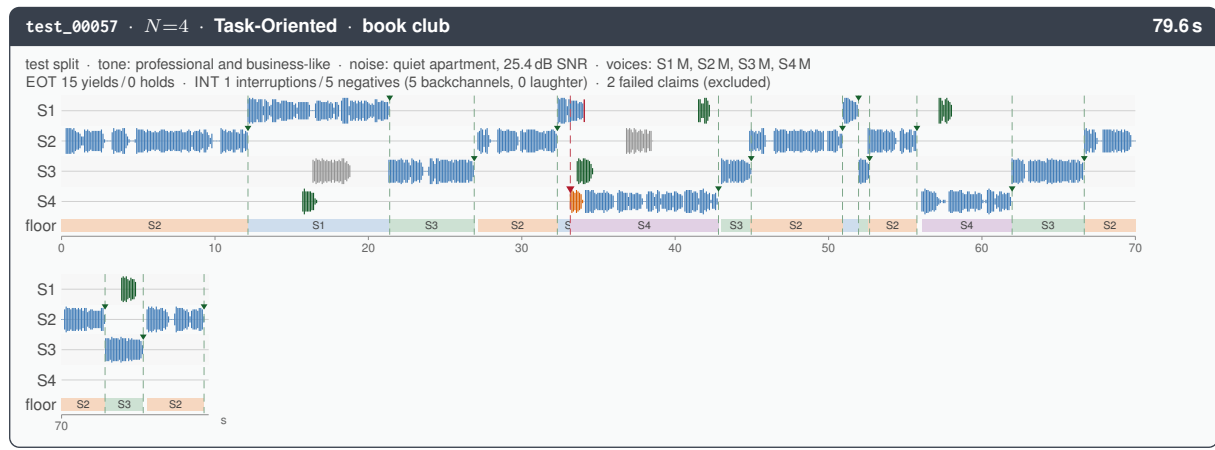
92.1–97.2 **S3** [turn] Okay but... aside from lighting, do we want open shelves anywhere? Or is everyone anti-dust? $\triangleright$ yields to S1

97.5–103.9 **S1** [turn] I'm 'pro put everything behind a door' but it's your dust allergy funeral. $\triangleright$ yields to S2

101.3–101.8 **S4** [backchannel] No way.

103.8–107.3 **S2** [turn] Honestly, yeah, I'm lazy, I'd never keep open shelves looking good. $\triangleright$ yields to S3

107.2–111.3 **S3** [turn] Haha, okay, closed cabinets win. I'll update that in the mockup file. $\triangleright$ yields, end



Transcript of test_00057

start–end (s) speaker [ground-truth label] text $\triangleright$ outcome

0.3–12.2 **S2** [turn] Alright, everyone, thanks for being on time. Let's get started. Our main agenda is to decide on next month's book and set up the discussion questions.
Anyone want to jump in first? $\triangleright$ yields to S1

12.2–18.1 **S1** [turn] Oh, can I just say, I finished 'Klara and the Sun' last night and it blew me away! Maybe we could do something with that?

15.7–16.7 **S4** [backchannel] Great choice.

16.4–18.8 **S3** [failed claim] Interesting, I was thinking the same thing actually—

18.2–21.4 **S1** [turn] —the AI themes are super timely, especially right now. $\triangleright$ yields to S3

21.3–26.9 **S3** [turn] Yeah, it gets at a lot of, um, ethical stuff. Plus, Kazuo Ishiguro just won another prize, didn't he? $\triangleright$ yields to S2

27.1–32.3 **S2** [turn] He did, yeah. So, do we have agreement, or do we want alternatives before we settle? $\triangleright$ yields to S1

32.4–34.1 **S1** [turn] I would vote for it, honestly. $\triangleright$ cut off

33.2–34.1 **S4** [interruption] Hold on— $\triangleright$ takes floor from S1

33.6–34.6 **S3** [backchannel] Works for me.

34.2–40.5 **S4** [turn] Let's not rush. I had one suggestion. Has anyone read 'The Nickel Boys'? It's short, powerful, and has lots to discuss.

36.8–38.5 **S2** [failed claim] I haven't read it, but I've heard it's—

40.6–42.8 **S4** [turn] —not a long read, that's why I picked it. $\triangleright$ yields to S3

41.5–42.3 **S1** [backchannel] Good call.

43.0–45.0 **S3** [turn] I'd be okay with either, honestly. $\triangleright$ yields to S2

44.9–50.9 **S2** [turn] Maybe we can vote? Just a quick show of hands for 'Klara and the Sun' versus 'The Nickel Boys'? $\triangleright$ yields to S1

51.0–52.0 **S1** [turn] Klara for me. $\triangleright$ yields to S3

52.0–52.7 **S3** [turn] Klara. $\triangleright$ yields to S2

52.6–55.8 **S2** [turn] I'll go with 'Klara and the Sun' too. That's three votes. $\triangleright$ yields to S4

56.1–57.7 **S4** [turn] Alright, fair enough.

57.2–58.1 **S1** [backchannel] Thanks, Tony.

57.9–62.0 **S4** [turn] So, who's on discussion questions this time? Last month was Jim, I think. $\triangleright$ yields to S3

62.0–66.7 **S3** [turn] I can draft some. Maybe three main ones, plus a backup or two in case the convo stalls. $\triangleright$ yields to S2

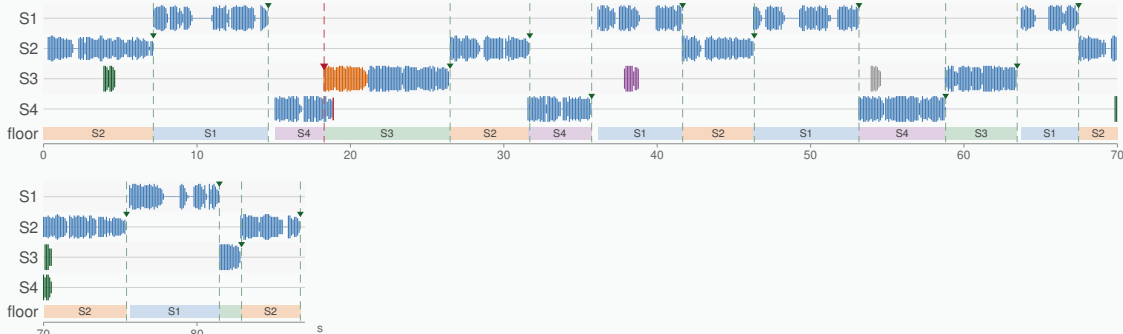
66.7–72.8 **S2** [turn] Perfect, Jim. Can you send them out by next Monday? That'll give us time to prep before we meet again. $\triangleright$ yields to S3

72.9–75.3 **S3** [turn] Yeah—I'll get them to everyone by Monday morning. $\triangleright$ yields to S2

73.9–74.9 **S1** [backchannel] Awesome, thanks!

75.6–79.3 **S2** [turn] Great, that covers everything. Thanks, everyone. Meeting adjourned. $\triangleright$ yields, end

test split · tone: curious and inquisitive · noise: quiet office, 27.1 dB SNR; alarm beep · voices: S1 F, S2 F, S3 F, S4 M
EOT 15 yields / 0 holds · INT 1 interruptions / 4 negatives (3 backchannels, 1 laughter) · 1 failed claims (excluded)

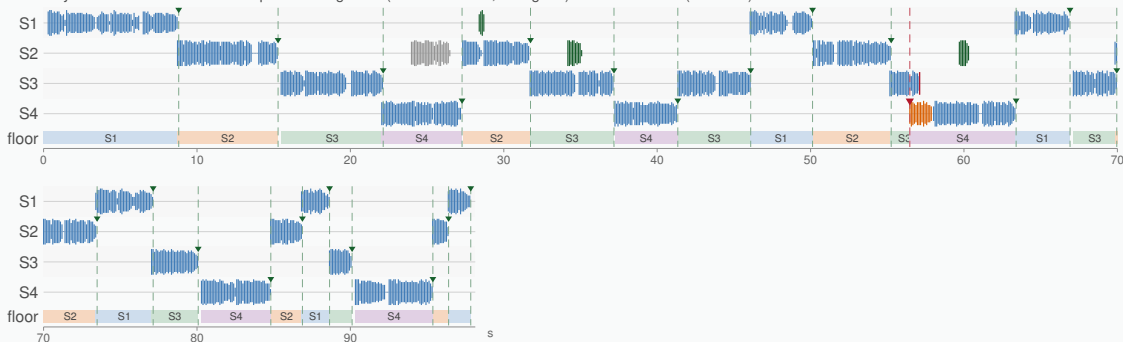


Transcript of test_00060

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–7.2 **S2** [turn] Alright, I'm opening the calendar. Let's just finalize who's driving to Nana's on Saturday and who's got the soccer drop-offs. ▷ yields to S1
3.9–4.7 **S3** [backchannel] Sounds good!
7.2–14.7 **S1** [turn] I volunteer as tribute for Nana. She'll want to critique my parallel parking again anyway. ▷ yields to S4
15.1–18.9 **S4** [turn] Um, I can take soccer, I think. Uh, unless you want me to swap, or— ▷ cut off
18.3–21.1 **S3** [interruption] Wait, wait, which team is it, the Tigers or the Flamingos? ▷ takes floor from S4
21.2–26.5 **S3** [turn] Because the Tigers have an away game at, like, 9 a.m. but the Flamingos don't play until one. ▷ yields to S2
26.5–31.7 **S2** [turn] Good catch, it's the Tigers this week. So that's the early slot. Is that too early for you, Dave? ▷ yields to S4
31.6–35.7 **S4** [turn] No, I mean, I'll set five alarms. I just probably need directions again. ▷ yields to S1
36.1–41.7 **S1** [turn] If you get lost at the same park again, you get a participation ribbon. ▷ yields to S2
37.9–38.9 **S3** [laughter] Haha, exactly.
41.6–46.3 **S2** [turn] I'll text the address. So, Jen, are you taking Nana both ways or need a pickup? ▷ yields to S1
46.3–53.2 **S1** [turn] I mean, if Nana's baking, I'll definitely need a ride home after the sugar crash. ▷ yields to S4
53.2–54.8 **S4** [turn] I can pick up, that's totally fine.
54.0–54.6 **S3** [failed claim] Oh! Wait—
54.9–58.8 **S4** [turn] —but only if someone reminds me by, like, texting, because I will forget. ▷ yields to S3
58.8–63.5 **S3** [turn] I can text! Just send me the actual time, otherwise I will also probably forget. ▷ yields to S1
63.7–67.5 **S1** [turn] So basically, we need a calendar for the calendar. ▷ yields to S2
67.5–75.4 **S2** [turn] Welcome to family logistics. Alright, to recap, Jen drives Nana, Dave does early Tigers drop-off, Lisa'll help with all reminders. ▷ yields to S1
69.9–70.6 **S4** [backchannel] Sounds good.
70.1–70.6 **S3** [backchannel] Yup!
75.6–81.5 **S1** [turn] If anyone else wants to swap with me, now is your moment of glory. ▷ yields to S3
81.5–82.9 **S3** [turn] No way, Nana's all yours. ▷ yields to S2
82.9–86.7 **S2** [turn] Okay, decision made. I'll follow up with the times. Thanks, team. ▷ yields, end

test split · tone: thoughtful and contemplative · noise: quiet apartment, 21.0 dB SNR; paper shuffle, dishes clinking · voices: S1 F, S2 F, S3 M, S4 F
EOT 23 yields / 0 holds · INT 1 interruptions / 3 negatives (3 backchannels, 0 laughter) · 1 failed claims (excluded)

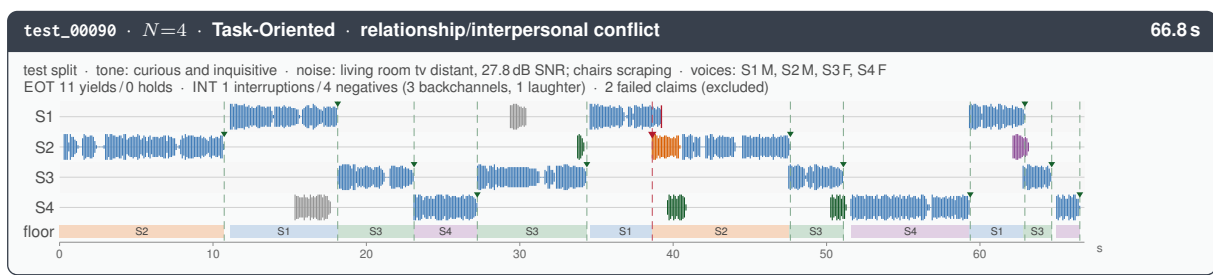


Transcript of test_00085

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.8 **S1** [turn] Alright, let's start. We're focusing on the onboarding flow today. Specifically, feedback on the preview screen updates. Quick round—who's had a chance to try the prototype? ▷ yields to S2
8.7–15.3 **S2** [turn] Yeah I tried it last night, and honestly, I really liked the new animation for the progress bar. It feels a lot smoother. ▷ yields to S3
15.5–22.1 **S3** [turn] I uh, I ran through it once, but I was a little confused by the copy under the checklist. Like, should it be more, um, explanatory? ▷ yields to S4
22.0–25.4 **S4** [turn] Can you clarify confused how? Was it just the text or the placement too?
24.0–26.5 **S2** [failed claim] Oh, same for me! I thought the checklist felt kind of—
25.5–27.3 **S4** [turn] —easy to miss? Or something else. ▷ yields to S2
27.3–31.7 **S2** [turn] Yeah, easy to miss. It just kind of blends in if you're not looking straight at it, you know? ▷ yields to S3
28.4–28.7 **S1** [backchannel] Mm.
31.7–37.2 **S3** [turn] Right, right, I think if we gave the checklist a little color block—like a shaded background—it'd pop more. ▷ yields to S4
34.2–35.1 **S2** [backchannel] Yeah, exactly.
37.2–41.3 **S4** [turn] Okay, so color block. What about the copy? Do we need to just add a little more context? ▷ yields to S3
41.4–46.1 **S3** [turn] I think so, yeah. Like instead of just 'Checklist' maybe call out what the user needs to do next. ▷ yields to S1

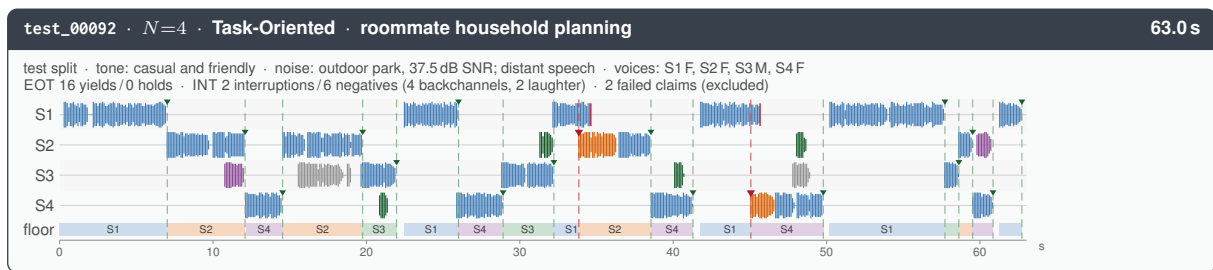
46.1–50.1 **S1** [turn] For example, ‘Complete these steps to get started.’ Something more actionable. ▷ yields to S2
 50.2–55.3 **S2** [turn] Yep, that’s much clearer. I can write up some alt copy options after the meeting if that helps. ▷ yields to S3
 55.2–57.1 **S3** [turn] Oh, that’d be... that’d be great. Thanks. ▷ cut off
 56.5–57.9 **S4** [interruption] Can we pause a second— ▷ takes floor from S3
 58.0–63.4 **S4** [turn] Sorry, but before we move on, did anyone notice lag on mobile? Mine felt a bit sluggish switching tabs. ▷ yields to S1
 59.7–60.3 **S2** [backchannel] Oh, uh-huh.
 63.3–66.9 **S1** [turn] Yes, same here. It hesitated when loading the resources section. ▷ yields to S3
 67.1–70.0 **S3** [turn] I only tested desktop, so I can’t speak to that. ▷ yields to S2
 69.8–73.5 **S2** [turn] I saw a little stutter, yeah. Like, not deal-breaking but enough to notice. ▷ yields to S1
 73.4–77.1 **S1** [turn] Should we file a ticket for that, or—actually, Mark, could you profile it later? ▷ yields to S3
 77.0–80.1 **S3** [turn] Yeah, sure, I’ll run some traces on Android after this. ▷ yields to S4
 80.3–84.8 **S4** [turn] Great. And can we lock in the animation speed? I think it’s good, but does anyone want it changed? ▷ yields to S2
 84.8–86.9 **S2** [turn] No, I think it’s just about perfect now. ▷ yields to S1
 86.8–88.6 **S1** [turn] Agreed. Let’s keep it as is. ▷ yields to S3
 88.6–90.1 **S3** [turn] No complaints from me. ▷ yields to S4
 90.3–95.4 **S4** [turn] Alright, then I’ll flag those two tweaks. Anything else on onboarding before we move to the next item? ▷ yields to S2
 95.4–96.4 **S2** [turn] Nope, not from me. ▷ yields to S1
 96.4–97.8 **S1** [turn] No, that covers it. ▷ yields, end



Transcript of test_00090

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–10.7 **S2** [turn] Alright everyone, thanks for joining. Our goal is to help Mark and Jess talk through that disagreement from last Friday and figure out some ways to move forward, okay? ▷ yields to S1
 11.1–18.1 **S1** [turn] Yeah, totally, uh—I just want to make sure Jess knows I, like, didn’t mean to shut her down in the meeting. I just get excited and... maybe I get loud. ▷ yields to S3
 15.3–17.7 **S4** [failed claim] I appreciate you starting with that, Mark.
 18.1–23.1 **S3** [turn] Can we clarify exactly what moment you’re referring to? Because there was a lot going on. ▷ yields to S4
 23.1–27.2 **S4** [turn] Right, Jess, you looked upset when Mark suggested the new plan for the rollout. ▷ yields to S3
 27.2–32.3 **S3** [turn] Well, I was. He interrupted while I was explaining why the old timeline mattered. I felt—
 29.4–30.5 **S1** [failed claim] I know, I know, but—
 32.4–34.4 **S3** [turn] —like my side didn’t get heard at all. ▷ yields to S1
 33.8–34.2 **S2** [backchannel] Mhm.
 34.6–39.2 **S1** [turn] Yeah, so, Jess, I really do want to hear your side. I just, uh—I thought we were behind and panicked. Can we try, like, letting each person finish next time before jumping in? ▷ cut off
 38.6–40.5 **S2** [interruption] Sorry, can I jump in real quick— ▷ takes floor from S1
 39.6–40.9 **S4** [backchannel] That sounds fair.
 40.6–47.6 **S2** [turn] Just to check, Jess, would it help if we had some kind of, like, hand signal or something for when someone wants to add on? ▷ yields to S3
 47.5–51.1 **S3** [turn] Honestly, that might. I just—I’d like to finish my thought next time. ▷ yields to S4
 50.3–51.3 **S4** [backchannel] Makes sense.
 51.6–59.4 **S4** [turn] So, if I’m hearing both of you, Mark’s worried about time and Jess wants to make sure she’s heard. Could we maybe agree to pause before responding? ▷ yields to S1
 59.3–62.9 **S1** [turn] Yeah, totally. Honestly, I could use five seconds to think sometimes. ▷ yields to S3
 62.2–63.2 **S2** [laughter] Haha, same here.
 62.8–64.7 **S3** [turn] I’m good with trying that, yeah. ▷ yields to S4
 65.0–66.5 **S4** [turn] Great. Thanks, everyone. ▷ yields, end

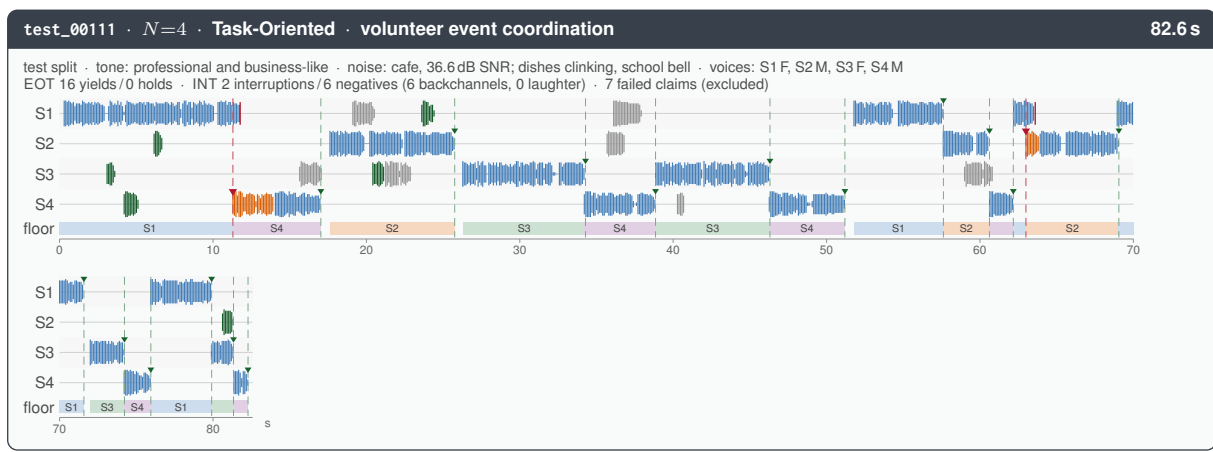


Transcript of test_00092

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–7.0 **S1** [turn] Alright, so, let’s just get started. We need to figure out who’s doing what for cleaning this week and also, um, groceries before Friday. ▷ yields to S2
 7.0–12.1 **S2** [turn] Right, I can do groceries if someone comes with me. Last time I, like, totally forgot the milk. ▷ yields to S4
 10.8–12.1 **S3** [laughter] Haha, yeah, I remember that.
 12.1–14.5 **S4** [turn] I can go with you, Kelly, that’s not a problem. ▷ yields to S2
 14.6–16.0 **S2** [turn] Yes! Awesome, thanks Tara.
 15.6–19.0 **S3** [failed claim] So uh, does that mean I’m off the hook for groceries this week, or—

16.2–19.8 **S2** [turn] —yeah, you can relax, Brian, unless you wanna carry like, the heavy stuff. ▷ yields to S3
 19.7–22.0 **S3** [turn] Nice, alright. But I'll do it next week—I promise. ▷ yields to S1
 20.9–21.4 **S4** [backchannel] Mm-hm.
 22.5–26.0 **S1** [turn] Okay, so, groceries sorted. What about cleaning? The kitchen's looking rough. ▷ yields to S4
 25.9–28.9 **S4** [turn] I vacuumed yesterday, but I didn't get to the counters, sorry. ▷ yields to S3
 28.8–32.2 **S3** [turn] Uh, yeah, I can do dishes tonight. I, uh, made the mess after all. ▷ yields to S1
 31.3–32.2 **S2** [backchannel] Thanks Brian!
 32.2–34.6 **S1** [turn] And I'll wipe down the counters and take out the trash tomorrow morning. ▷ cut off
 33.9–36.3 **S2** [interruption] Wait, do we have enough trash bags? I think we're low. ▷ takes floor from S1
 36.5–38.6 **S2** [turn] 'Cause last time I checked, there were, like, two left. ▷ yields to S4
 38.6–41.3 **S4** [turn] Oh, I can add that to our grocery list, no worries. ▷ yields to S1
 40.1–40.7 **S3** [backchannel] Cool, cool.
 41.8–45.7 **S1** [turn] One more thing—can someone do the bathroom? I did it last weekend and it's seriously— ▷ cut off
 45.1–46.6 **S4** [interruption] I'll take care of it, Alison. ▷ takes floor from S1
 46.7–49.8 **S4** [turn] Honestly, I don't mind. I usually do it Fridays anyway. ▷ yields to S1
 47.8–48.9 **S3** [failed claim] Seriously, you're a legend.
 48.1–48.7 **S2** [backchannel] You're the best.
 50.2–57.7 **S1** [turn] Alright, so just to check, groceries are Kelly and Tara, Brian does dishes, I've got the counters and trash, and Tara's on the bathroom. Everyone good? ▷ yields to S3
 57.7–58.6 **S3** [turn] Yup, works for me. ▷ yields to S2
 58.6–59.5 **S2** [turn] Sounds perfect! ▷ yields to S4
 59.5–60.9 **S4** [turn] Yeah, looks good to me. ▷ yields to S1
 59.8–60.7 **S2** [laughter] Sweet, team work!
 61.3–62.7 **S1** [turn] Alright, meeting adjourned. ▷ yields, end

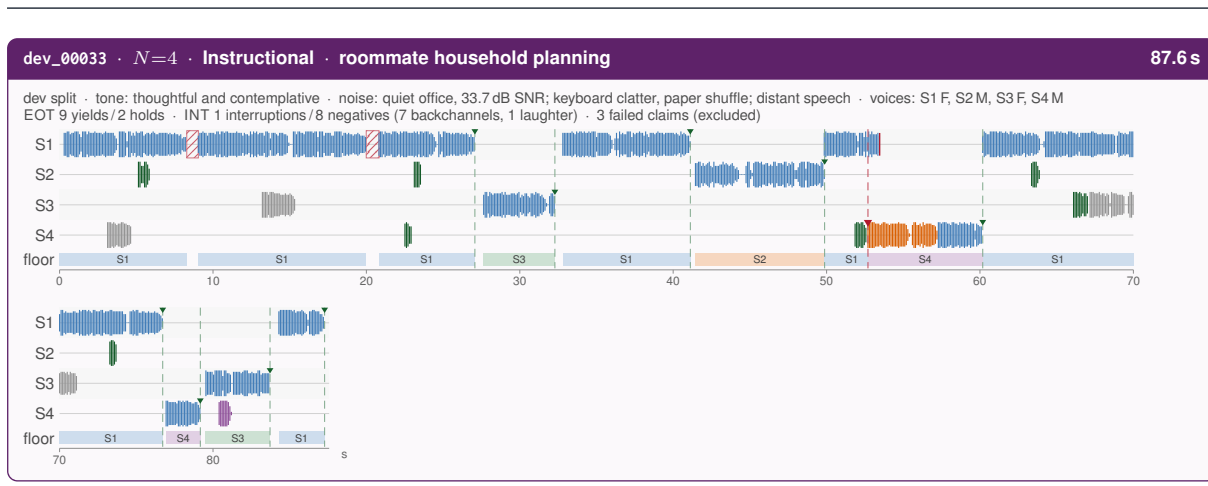


Transcript of test_00111

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.0 **S1** [turn] Thanks everyone for coming together on such short notice. Let's just jump in—our goal is to finalize Saturday's volunteer event details today.
 3.1–3.6 **S3** [backchannel] Mm-hm.
 4.2–5.2 **S4** [backchannel] Yeah, I'm ready.
 6.2–6.7 **S2** [backchannel] Right.
 8.0–11.8 **S1** [turn] Great, so, first, check-in volunteers. We currently have six, but two people haven't confirmed— ▷ cut off
 11.3–13.9 **S4** [interruption] Do we have backups for check-in if those two drop out? ▷ takes floor from S1
 14.1–17.0 **S4** [turn] I can reach out to a couple folks from last month if needed. ▷ yields to S2
 15.7–17.1 **S3** [failed claim] That might be smart.
 17.6–22.3 **S2** [turn] I could also cover check-in for the first hour. My other duties don't ramp up until later.
 19.1–20.6 **S1** [failed claim] That's really helpful, Mark.
 20.4–21.2 **S3** [backchannel] Awesome.
 21.3–23.0 **S3** [failed claim] Um, about the refreshments—
 22.5–25.8 **S2** [turn] —And I'll bring my name badge labels, so no worries if we run out. ▷ yields to S3
 23.6–24.4 **S1** [backchannel] Perfect.
 26.3–34.3 **S3** [turn] Okay, so for refreshments, Angela confirmed juice and water, but I haven't heard back from the bakery yet. I'll call again this afternoon. ▷ yields to S4
 34.2–38.8 **S4** [turn] If they can't do it, could we switch to bagels from Corner Cafe? I know they open early. ▷ yields to S3
 35.7–36.9 **S2** [failed claim] Yeah, I like that.
 36.1–38.0 **S1** [failed claim] That's a backup option, for sure.
 38.9–41.9 **S3** [turn] I'll update everyone by tonight, one way or the other.
 40.3–40.8 **S4** [failed claim] Can I—
 42.0–46.3 **S3** [turn] —And if people have dietary needs, just email me. I'll handle the labels. ▷ yields to S4
 46.3–51.2 **S4** [turn] Sorry! I was just gonna ask about the supply pickup—do we need another car for the folding tables? ▷ yields to S1
 51.8–57.6 **S1** [turn] Actually, I can fit the tables and signs in my trunk, but if we end up adding chairs, I might need someone else to drive. ▷ yields to S2
 57.6–60.6 **S2** [turn] I'll bring my SUV just in case. It's no trouble. ▷ yields to S4
 59.0–60.9 **S3** [failed claim] That takes a load off, thanks.
 60.6–62.2 **S4** [turn] Great teamwork, everyone! ▷ yields to S1
 62.2–63.6 **S1** [turn] Okay, last item— ▷ cut off
 63.0–63.8 **S2** [interruption] Oh, wait— ▷ takes floor from S1
 63.9–69.1 **S2** [turn] Sorry, just one question. Do we have all the liability waivers printed, or should I take care of those? ▷ yields to S1
 68.9–71.6 **S1** [turn] I'll print them this evening and bring extras, don't worry. ▷ yields to S3
 72.0–74.2 **S3** [turn] Thanks for organizing all this, really. ▷ yields to S4
 74.2–75.9 **S4** [turn] Yeah, we appreciate you, Sam! ▷ yields to S1
 75.9–79.9 **S1** [turn] Aw, thank you. Okay, I'll send an updated checklist after this, and we're good to go! ▷ yields to S3

79.9–81.3 **S3** [turn] See you all Saturday. ▷ yields to S4
 80.6–81.3 **S2** [backchannel] See you.
 81.4–82.3 **S4** [turn] Alright! Bye! ▷ yields, end

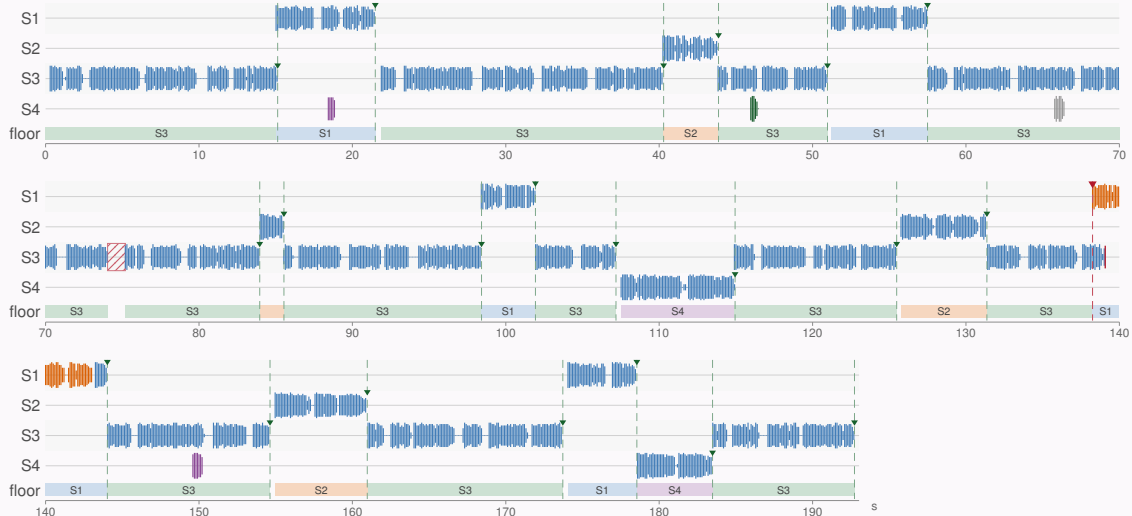


Transcript of dev_00033

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.3 **S1** [turn] Alright, so the new cleaning schedule I made is in the kitchen, but I thought it'd be good to walk through it together so we're all on the same page. ▷ holds, pauses
 3.1–4.7 **S4** [failed claim] Yeah, thanks for organizing it.
 5.1–5.9 **S2** [backchannel] Makes sense.
 9.0–20.0 **S1** [turn] So basically, instead of rotating every Sunday, each person has one task for the entire week. For example, someone takes out trash all week, someone else does dishes, and so on. ▷ holds, pauses
 13.2–15.3 **S3** [failed claim] But what if someone's away a couple days—
 20.8–27.1 **S1** [turn] —right, if that's the case, just let us know, we can swap for that week. It's flexible, as long as we communicate. ▷ yields to S3
 22.5–23.0 **S4** [backchannel] Mm-hm.
 23.1–23.6 **S2** [backchannel] Got it.
 27.6–32.3 **S3** [turn] So... how strict are we about actually doing the job each day? What if someone's just... late? ▷ yields to S1
 32.8–41.1 **S1** [turn] Ideally it gets done that day, but we're not policing anyone. If dishes pile up for two days, just please communicate so nobody gets frustrated. ▷ yields to S2
 41.5–49.9 **S2** [turn] Would it help if we just text when something slips? Like, if I know I'm swamped and can't do the trash for a night, should I just drop it in the group chat? ▷ yields to S1
 49.8–53.5 **S1** [turn] Yeah—that's exactly what I had in mind. Just a quick heads-up is all I'm asking. ▷ cut off
 51.9–52.6 **S4** [backchannel] That's fair.
 52.7–57.2 **S4** [interruption] Can I ask, uh, does the schedule include like, the living room? Or is it just kitchen and bathroom? ▷ takes floor from S1
 57.3–60.2 **S4** [turn] Because honestly, that's where stuff tends to pile up for us. ▷ yields to S1
 60.2–68.9 **S1** [turn] Yeah, so living room is a task as well, just like kitchen. Whoever gets that one is responsible for picking up and vacuuming by the end of the week.
 63.4–63.9 **S2** [backchannel] Cool.
 66.1–67.1 **S3** [backchannel] Ah, okay.
 67.2–71.2 **S3** [failed claim] But how do we make sure it's, like, fair over time? Some chores are worse than others—
 69.0–76.7 **S1** [turn] —yeah, that's why it rotates each week, so nobody gets stuck with bathroom more than anyone else. I kept a list to make sure it evens out. ▷ yields to S4
 73.3–73.8 **S2** [backchannel] Nice.
 76.9–79.2 **S4** [turn] Honestly, this sounds way better than what we had. ▷ yields to S3
 79.5–83.7 **S3** [turn] We'll see, I reserve judgment until we've survived your week with bathroom duty. ▷ yields to S1
 80.4–81.2 **S4** [laughter] Haha, right.
 84.3–87.3 **S1** [turn] Deal. But if it doesn't work, we can tweak it. I'm not attached. ▷ yields, end

dev split · tone: enthusiastic and energetic · noise: living room tv distant, 31.6 dB SNR; notification chime · voices: S1 F, S2 M, S3 F, S4 M
EOT 21 yields / 1 holds · INT 1 interruptions / 3 negatives (1 backchannels, 2 laughter) · 1 failed claims (excluded)

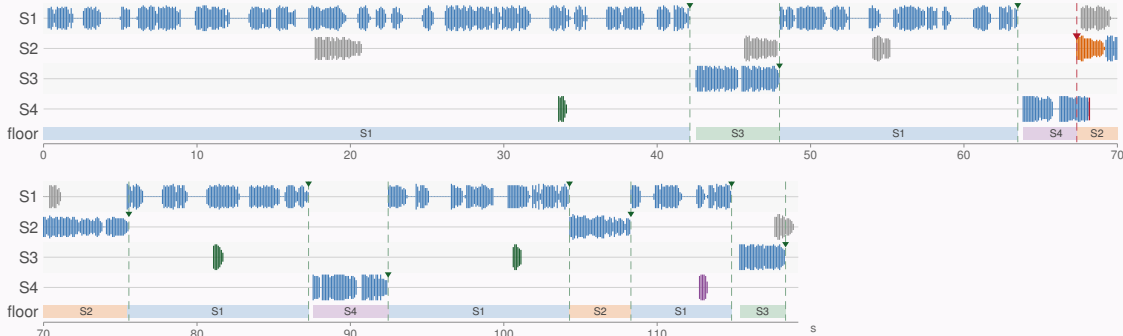


Transcript of dev_00036

start—end (s) speaker [ground-truth label] text ▷ outcome

- 0.3–15.1 **S3** [turn] Alright everyone, thanks for coming today! So we're going to break down the main steps for planning a home renovation, from initial brainstorming all the way through hiring contractors. I'll walk you through the process, but seriously, ask questions anytime, okay? ▷ yields to S1
- 15.0–21.5 **S1** [turn] Awesome, I've been kinda dying to figure this out. Where do you even start, like, do you just make a Pinterest board or something? ▷ yields to S3
- 18.4–18.9 **S4** [laughter] haha
- 21.9–40.3 **S3** [turn] Not a bad idea! Actually, the first thing I always recommend is to get really clear about why you want to renovate. Are you looking for more space, better function, changing the style? Once you know that, then yes, start collecting inspiration—Pinterest is great, magazines, even just photos you take in cafes you like. ▷ yields to S2
- 40.2–43.9 **S2** [turn] So you're saying define the problem before you jump into solutions. ▷ yields to S3
- 43.8–51.0 **S3** [turn] Exactly. Because if you just start picking tile, you might end up with something gorgeous but totally wrong for how you actually live. ▷ yields to S1
- 46.0–46.5 **S4** [backchannel] Right
- 51.2–57.5 **S1** [turn] Okay, I get that. But then how do you figure out like, what's realistic for your budget? I always get lost in that part. ▷ yields to S3
- 57.5–74.0 **S3** [turn] Yeah, budget's huge. The next step once you have your goals is to do some research about costs. Look up average renovation costs in your area for the kind of work you want—websites like HomeAdvisor or talking to neighbors can help. And always add at least ten percent for surprises. ▷ holds, pauses
- 65.8–66.5 **S4** [failed claim] But ten—
- 75.2–84.0 **S3** [turn] —ten percent, minimum! Sometimes twenty if it's an old house. Trust me, there's always some weird plumbing or electrical thing hiding behind the walls. ▷ yields to S2
- 84.0–85.5 **S2** [turn] Oh, you have stories, huh? ▷ yields to S3
- 85.5–98.4 **S3** [turn] So many. But let's keep you all from collecting your own horror stories. Once you know your priorities and rough budget, start sketching out your ideas. Even stick figures will help a designer or contractor understand what you want. ▷ yields to S1
- 98.4–101.9 **S1** [turn] Like, literal stick people? I feel like mine just look like blobs. ▷ yields to S3
- 101.9–107.2 **S3** [turn] Haha, blobs work as long as you can explain them! It's about communication, not artistry. ▷ yields to S4
- 107.5–114.9 **S4** [turn] So what about timing? Like, how long should you expect a kitchen redo to actually take? Because my cousin thought hers would be a month and it was four, I swear. ▷ yields to S3
- 114.9–125.5 **S3** [turn] Oh, that's the classic trap! Always double the time you hope for, especially for kitchens. There's demo, waiting for permits, ordering cabinets—if one thing's delayed, the whole thing backs up. ▷ yields to S2
- 125.8–131.4 **S2** [turn] So when do you actually bring in professionals? Do you need an architect or just a general contractor—or both? ▷ yields to S3
- 131.4–139.1 **S3** [turn] That depends on how big the project is. For anything structural or if you're moving walls or plumbing, definitely get an architect or a structural engineer involved. For smaller stuff, a good contractor can sometimes draw up simple plans themselves. ▷ cut off
- 138.2–143.1 **S1** [interruption] Wait—so if you're like, just painting or swapping cabinets, you don't need permits or anything? ▷ takes floor from S3
- 143.2–144.0 **S1** [turn] Or do you? ▷ yields to S3
- 144.0–154.6 **S3** [turn] Usually for painting, you're fine. Cabinets too, as long as you're not changing the layout or moving plumbing or electric. But: always check your city's permit rules. They love surprises. ▷ yields to S2
- 149.6–150.3 **S4** [laughter] haha right
- 155.0–161.0 **S2** [turn] What's the best way to actually pick a contractor? All those reviews online, it's just overwhelming sometimes. ▷ yields to S3
- 161.0–173.7 **S3** [turn] That's a great question. Word of mouth is gold—talk to friends or neighbors who've done renovations. Get at least three quotes. And always ask to see recent examples of their work—don't just trust the photos on their website. ▷ yields to S1
- 174.1–178.5 **S1** [turn] Can I just say this is making me actually want to start? It sounds less scary somehow. ▷ yields to S4
- 178.6–183.5 **S4** [turn] Yeah, for real. I thought it would be way more complicated, but it's nice seeing it all broken down like this. ▷ yields to S3
- 183.5–192.7 **S3** [turn] I love hearing that! Just remember, take it step by step. If you hit a snag, just ask for advice—you've got plenty of resources and, well, you've got me in your corner now. ▷ yields, end

test split · tone: curious and inquisitive · noise: office hvac, 26.2 dB SNR; paper shuffle · voices: S1 F, S2 M, S3 M, S4 M
EOT 10 yields / 0 holds · INT 1 interruptions / 4 negatives (3 backchannels, 1 laughter) · 5 failed claims (excluded)



Transcript of test_00011

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–29.3 **S1** [turn] Alright, so for this round, it's the picture trivia. Basically, they'll show us a close-up of some random object, and we have to guess what it actually is. Sometimes it's, like, painfully obvious and other times it's a macro shot of, I don't know, a kiwi or something. So yeah, you just write down your best guess on the answer sheet.

17.7–20.8 **S2** [failed claim] Oh, cool! Do we get like, uh, points for partial answers, or—

29.4–42.1 **S1** [turn] —No, it's extremely unforgiving. You either get it right or you don't. If it's a kiwi and you say cucumber, you get nothing but disappointment. ▷ yields to S3

33.6–34.1 **S4** [backchannel] uh-huh

42.5–48.0 **S3** [turn] So, um, do they give us any hints if nobody gets it, or is it just permanent mystery if we all fail? ▷ yields to S1

45.7–47.9 **S2** [failed claim] Yeah, like do they ever just show us what it was?

48.0–63.5 **S1** [turn] No, luckily they usually put us out of our misery at the end of the round. They'll be like, 'That weird green blob was, in fact, a kiwi.' So you get closure, just not points. ▷ yields to S4

54.1–55.3 **S2** [failed claim] Ah, got it, got it.

63.9–68.2 **S4** [turn] Is there like a time limit for each photo, or can we argue for three minutes about whether it's a tennis ball or a lime? ▷ cut off

67.4–69.2 **S2** [interruption] Oh, like, a minute or something? ▷ takes floor from S1

67.6–71.1 **S1** [failed claim] There's definitely a time limit, but it's not—

69.3–75.6 **S2** [turn] Because last time I swear we all started debating the difference between a lemon and a squash and then the timer ran out on us. ▷ yields to S1

75.5–87.3 **S1** [turn] Yeah, exactly. It's like a minute per picture. If you're still bickering by the end, just put your wildest guess and move on with your life. ▷ yields to S4

81.1–81.7 **S3** [backchannel] right

87.6–92.5 **S4** [turn] I see. And do we have to agree on one answer as a team, or can we submit different guesses? ▷ yields to S1

92.5–104.3 **S1** [turn] Team consensus, unfortunately. Democracy usually leads to chaos, but only one answer gets written down per round for our team. ▷ yields to S2

100.6–101.2 **S3** [backchannel] mm-hm

104.3–108.3 **S2** [turn] Okay, so we just argue and then whoever grabs the pen first wins, got it. ▷ yields to S1

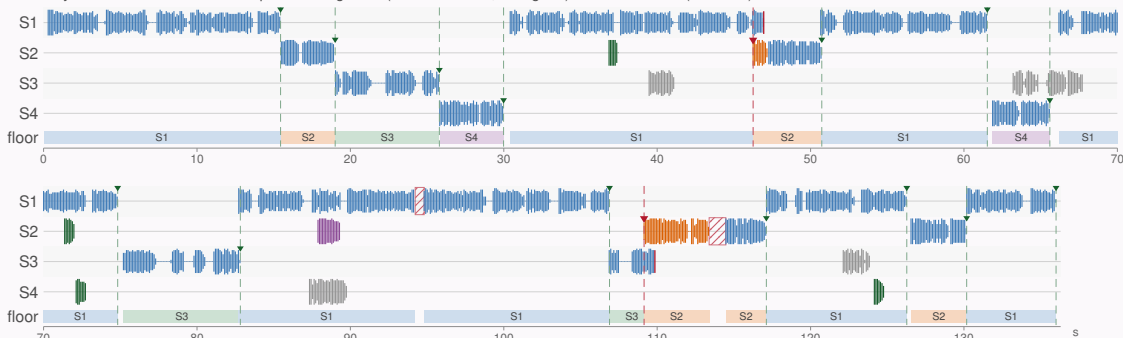
108.3–114.8 **S1** [turn] Exactly. May the fastest writer win. Or, you know, the most stubborn. ▷ yields to S3

112.8–113.3 **S4** [laughter] haha

115.4–118.4 **S3** [turn] Alright, I'll try not to freeze up when it's our turn to guess. ▷ yields, end

117.7–118.9 **S2** [failed claim] You got this, man.

test split · tone: enthusiastic and energetic · noise: office hvac, 24.2 dB SNR; laughter offmic · voices: S1 M, S2 F, S3 F, S4 M
EOT 14 yields / 2 holds · INT 2 interruptions / 5 negatives (4 backchannels, 1 laughter) · 4 failed claims (excluded)



Transcript of test_00016

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–15.4 **S1** [turn] Alright, so tonight we're playing Quest for the Cursed Chalice. I'll run through the basics and then we can get started. Each of you gets a character card, and your goal is to collect three relic tokens before anyone else, then make it back to the castle. ▷ yields to S2

15.5–19.0 **S2** [turn] Ooh, love it already. Do we choose our characters, or is it random? ▷ yields to S3

19.0–25.8 **S3** [turn] Yeah, because if I get the Bard again, I refuse responsibility for any ballads. ▷ yields to S4

25.8–30.0 **S4** [turn] Uh, sorry, so—do the relics, like, have to be different, or can you have three of the same? ▷ yields to S1

30.4–42.6 **S1** [turn] Good questions! You'll draft characters, so you get some choice. And yes, the relics do have to be different types. There's the Chalice Fragment, Ancient Coin, and Mystic Gem. You'll need one of each.

36.9–37.5 **S2** [backchannel] Gotcha.

39.5–41.2 **S3** [failed claim] That's... not ominous at all.

42.7–46.9 **S1** [turn] Now, on your turn, you get two actions. You can move, search a room for relics, or sabotage— ▷ cut off

46.3–47.2 **S2** [interruption] Wait, sabotage? ▷ takes floor from S1

47.3–50.7 **S2** [turn] Can you sabotage other people’s relics, or just like, slow them down? ▷ yields to S1

50.7–61.5 **S1** [turn] Both, technically! You can set traps in rooms to block someone, or if you’re sneaky, you can steal a relic if you’re in the same room as them. There’s a dice roll for that, though. ▷ yields to S4

61.9–65.6 **S4** [turn] Oh, um, so if you lose a relic, do you have to go get another of that type? ▷ yields to S1

63.2–67.8 **S3** [failed claim] I foresee an immediate and total collapse of alliances.

66.2–74.8 **S1** [turn] Yeah, exactly. You’re not eliminated or anything, but if someone steals or traps away your Mystic Gem, you’ll have to go find another one. ▷ yields to S3

71.4–72.1 **S2** [backchannel] Right, right.

72.1–72.8 **S4** [backchannel] Yikes.

75.2–82.8 **S3** [turn] If I’m in the same room as someone, can I just, I don’t know, jump out a window instead? ▷ yields to S1

82.7–94.2 **S1** [turn] Haha, you could—well, if you have the Acrobat card you actually can! There are some fun character powers. But otherwise, it’s just dice and sabotage and maybe a little running for your life. ▷ holds, pauses

87.3–89.8 **S4** [failed claim] Uh... does everyone get a special power then?

87.9–89.4 **S2** [laughter] Haha, I want the Acrobat now.

94.8–106.9 **S1** [turn] —Yeah, yeah, every character has a unique ability listed on their card. Some let you move extra spaces, some get bonuses searching. Acrobat can escape windows. You’ll see when we draft. ▷ yields to S3

106.9–109.9 **S3** [turn] Perfect. I call dibs on the weirdest one. ▷ cut off

109.1–113.4 **S2** [interruption] So—is there a way to block someone from getting into the castle? Like, camp the door? ▷ takes floor from S3 ▷ holds, pauses

114.5–117.1 **S2** [turn] Not that I would! I’m just, you know, strategizing early. ▷ yields to S1

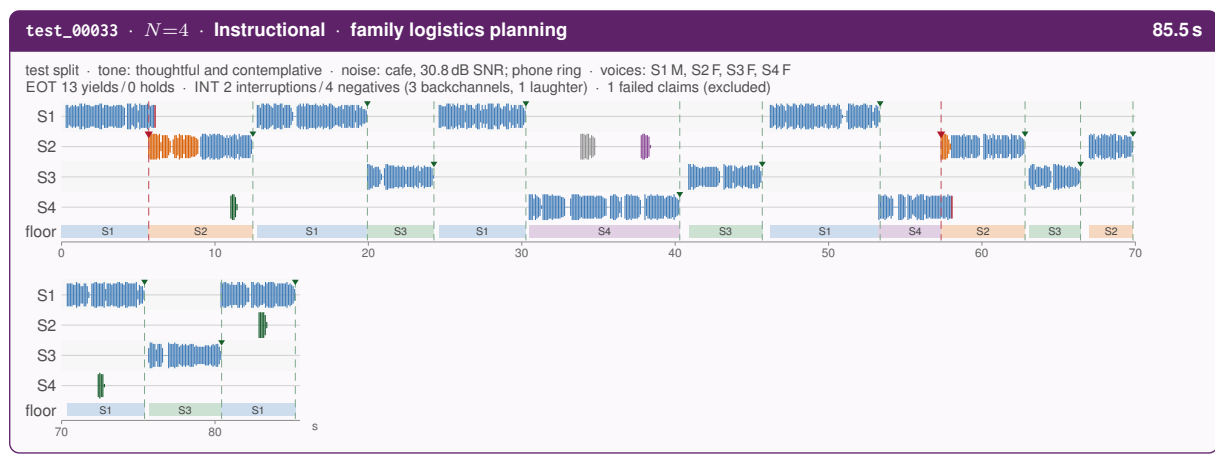
117.1–126.3 **S1** [turn] Uh-huh, totally innocent. Yes, you can lay a trap right at the castle entrance if you want. But there are only three traps, so use them wisely. ▷ yields to S2

122.1–123.9 **S3** [failed claim] I feel very safe right now.

124.2–124.8 **S4** [backchannel] Gulp.

126.5–130.2 **S2** [turn] Let’s go, I need to see these character cards. Ready to be chaotic. ▷ yields to S1

130.2–136.0 **S1** [turn] Alright, deal out three character cards each, and pick your favorite. Let the dungeon-crawling begin! ▷ yields, end



Transcript of test_00033

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.1 **S1** [turn] Um, okay, so I thought maybe we could walk through the weekend plans, uh, so everyone’s on the same page? There’s a lot going on with the kids’ activities and grandma coming over— ▷ cut off

5.7–9.0 **S2** [interruption] Oh, that’s so helpful, thanks! Can we start with Saturday morning? ▷ takes floor from S1

9.1–12.5 **S2** [turn] Because I think we had talked about maybe switching who does soccer drop-off? ▷ yields to S1

11.0–11.5 **S4** [backchannel] mm-hm

12.8–19.9 **S1** [turn] Yeah, so, uh, soccer starts at nine. If I take Maddie this time, then maybe, um, speaker_3 can pick up groceries? ▷ yields to S3

20.0–24.3 **S3** [turn] That works for me. Should I get everything on the list, or just what’s needed for Sunday dinner? ▷ yields to S1

24.6–30.3 **S1** [turn] Maybe, uh, just what’s needed for dinner, since I think we’ve got enough snacks and breakfast stuff already. ▷ yields to S4

30.5–36.6 **S4** [turn] Right, right, and I offered to mow the lawn, but does that make sense in the morning while you’re out, or should I wait?

33.9–34.9 **S2** [failed claim] Oh, good question—

36.7–40.3 **S4** [turn] —because it might be loud, if Grandma’s napping then, she gets fussy. ▷ yields to S3

37.8–38.4 **S2** [laughter] haha true

40.9–45.7 **S3** [turn] Maybe afternoon’s just safer for mowing. I can time my grocery run so I’m back by then. ▷ yields to S1

46.2–53.4 **S1** [turn] Um, yeah, that’s probably better. Then, after lunch, we need to get the guest room ready. Grandma usually arrives by three. ▷ yields to S4

53.3–58.0 **S4** [turn] I can do the sheets! And I’ll, uh, stock the bathroom with new towels and her favorite soap. ▷ cut off

57.3–57.9 **S2** [interruption] So, wait— ▷ takes floor from S4

58.0–62.8 **S2** [turn] Do we need to pick up her medication, or did she say she was bringing it herself this time? ▷ yields to S3

63.1–66.4 **S3** [turn] She texted me, she’s bringing it. But I’ll double-check, just in case. ▷ yields to S2

67.0–69.8 **S2** [turn] Perfect. And Sunday, there’s that art fair in the park, right? ▷ yields to S1

70.4–75.4 **S1** [turn] Yeah, it starts at ten. If everyone’s up for it, we can all go together after breakfast. ▷ yields to S3

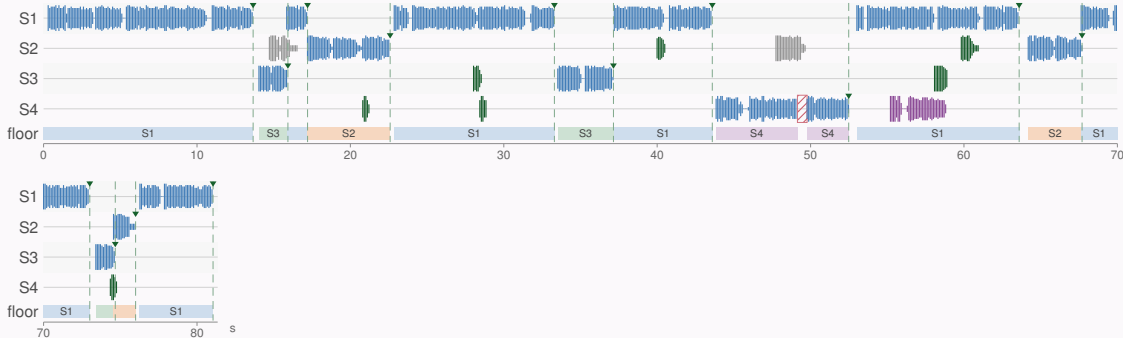
72.4–72.8 **S4** [backchannel] yay

75.7–80.4 **S3** [turn] Sounds like a plan. Thanks for getting us coordinated, I think this is the smoothest we’ve ever done. ▷ yields to S1

80.4–85.2 **S1** [turn] Haha, well, let’s see how it goes. I’ll, um, type this all up so we don’t forget anything. ▷ yields, end

82.9–83.4 **S2** [backchannel] awesome

test split · tone: casual and friendly · noise: quiet office, 24.8 dB SNR; door knock · voices: S1 F, S2 M, S3 M, S4 M
EOT 14 yields / 1 holds · INT 0 interruptions / 8 negatives (7 backchannels, 1 laughter) · 2 failed claims (excluded)



Transcript of test_00068

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–13.7 **S1** [turn] Okay, so I thought we could talk through how the new cleaning schedule will work, just so everyone's on the same page. I made a chart for the kitchen, bathroom, and living room, and I split up the chores so we each get two areas per week. That way, nobody's stuck with the worst stuff every time. ▷ yields to S3

14.0–15.9 **S3** [turn] Nice, that sounds pretty fair actually. ▷ yields to S1

14.7–16.6 **S2** [failed claim] Yeah, uh, quick question—

15.8–17.2 **S1** [turn] —oh, go ahead, sorry. ▷ yields to S2

17.2–22.6 **S2** [turn] No, I was just gonna ask, like, when does the week start for chores? Are we talking Sunday to Saturday or— ▷ yields to S1

20.8–21.2 **S4** [backchannel] mm-hm

22.9–33.3 **S1** [turn] Yeah, good question. I figured Monday mornings would be the starting point, so if you have, say, the bathroom that week, you'd want to have it done anytime before the next Monday rolls around. That way, it's easy to remember. ▷ yields to S3

28.0–28.6 **S3** [backchannel] uh-huh

28.4–28.9 **S4** [backchannel] right

33.5–37.2 **S3** [turn] So, if I'm out of town a weekend, can I swap my week with someone, or...? ▷ yields to S1

37.2–43.6 **S1** [turn] Exactly, just let everyone know in the group chat if you need to trade. I'm totally fine with people switching as long as we all agree. ▷ yields to S4

40.0–40.5 **S2** [backchannel] cool

43.8–49.1 **S4** [turn] So, um, what about supplies? Are we splitting costs or is that part of the rotation too? ▷ holds, pauses

47.7–49.7 **S2** [failed claim] Oh, like the paper towels and stuff?

49.8–52.5 **S4** [turn] Yeah, just, cleaning sprays, toilet paper, all that. ▷ yields to S1

53.0–63.6 **S1** [turn] Great point. I was thinking we could each pitch in five bucks a month and then I'll just Venmo whoever buys the stuff, or whoever gets it just posts a receipt and we pay them back. That way no one's, like, stuck buying everything. ▷ yields to S2

55.2–58.8 **S4** [laughter] haha, yeah, nobody wants to be the sole toilet paper hero.

58.1–59.0 **S3** [backchannel] yeah, that works

59.8–61.0 **S2** [backchannel] Makes sense.

64.2–67.7 **S2** [turn] Right? That was me last semester, I swear I bought, like, a truckload. ▷ yields to S1

67.6–73.0 **S1** [turn] Haha, yeah, no more heroes needed. So does everyone feel good about the schedule, or is there something I missed? ▷ yields to S3

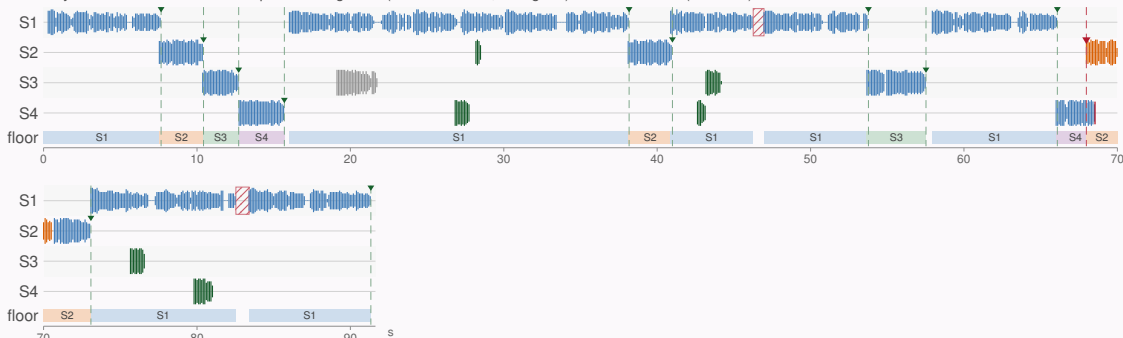
73.4–74.7 **S3** [turn] Nope, looks solid to me. ▷ yields to S2

74.4–74.8 **S4** [backchannel] yeah

74.6–76.0 **S2** [turn] All good over here. ▷ yields to S1

76.3–81.0 **S1** [turn] Awesome, thanks, everyone. I'll share the chart in the chat and if anyone wants tweaks, just let me know. ▷ yields, end

test split · tone: enthusiastic and energetic · noise: quiet apartment, 29.0 dB SNR; door knock · voices: S1 M, S2 M, S3 M, S4 M
EOT 11 yields / 2 holds · INT 1 interruptions / 6 negatives (6 backchannels, 0 laughter) · 1 failed claims (excluded)



Transcript of test_00082

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–7.7 **S1** [turn] Hey guys, thanks for hopping on! So today we're gonna walk through setting up remote desktop, step-by-step. It's easier than it sounds, I promise. ▷ yields to S2

7.5–10.4 **S2** [turn] Um, yeah, thanks, I'm—I'm really not sure where to start with it. ▷ yields to S3

10.4–12.7 **S3** [turn] Same here, I keep getting lost in the settings. ▷ yields to S4

12.7–15.7 **S4** [turn] Yeah, I... I couldn't even find the option last time I tried. ▷ yields to S1

16.0–24.0 **S1** [turn] No worries at all! So, first thing—on your main computer, click the Start menu and type 'remote desktop settings.' You should see it pop up—just click that.

19.1–21.7 **S3** [failed claim] Wait, is that the same as Remote Assistance, or—

24.1–30.0 **S1** [turn] —good question! Nope, it's Remote Desktop, not Assistance. The icons look super similar though, easy mix-up.

26.8–27.8 **S4** [backchannel] ohhh, okay

28.2–28.6 **S2** [backchannel] ahh

30.1–38.2 **S1** [turn] Alright, so now you should see a switch labeled 'Enable Remote Desktop.' Flip that on. You'll probably get a prompt about network access—just accept it. ▷ yields to S2

38.1–41.0 **S2** [turn] Do we, um, need to write down the computer name somewhere? ▷ yields to S1

40.9–46.3 **S1** [turn] Yes! Great catch, you'll want to note that name down, that's how you'll connect later, so maybe jot it or screenshot it. ▷ holds, pauses

42.6–43.2 **S4** [backchannel] mm-hm

43.2–44.2 **S3** [backchannel] Right, okay.

47.0–53.8 **S1** [turn] Now, on the other computer—the one you'll use to connect—open the Remote Desktop app. You just search 'Remote Desktop' in Start. Got it? ▷ yields to S3

53.7–57.5 **S3** [turn] Yeah, I've got it open, but it always just gives me a blank screen for some reason. ▷ yields to S1

57.9–66.1 **S1** [turn] That can happen if the firewall is blocking it, but let's get everyone to the 'add a new connection' screen first. Click 'Add,' then enter your other computer's name. ▷ yields to S4

66.0–68.5 **S4** [turn] Uh, is it... normal for it to ask for my credentials again? ▷ cut off

68.0–70.6 **S2** [interruption] Yeah, mine did that too, like, username and password? ▷ takes floor from S4

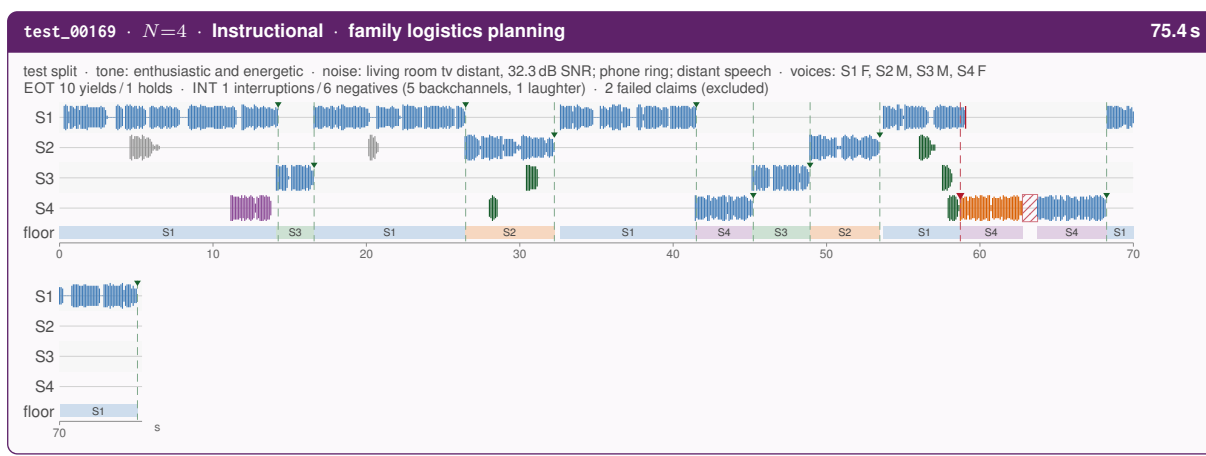
70.7–73.1 **S2** [turn] I just don't remember which username to use. ▷ yields to S1

73.1–82.5 **S1** [turn] Yep, exactly, it needs your Windows username and password for the main computer. It's usually your login name—if you're not sure, open a command line on the main machine and type 'whoami'. ▷ holds, pauses

75.7–76.6 **S3** [backchannel] oh, cool trick

79.8–81.0 **S4** [backchannel] yeah, super helpful

83.4–91.3 **S1** [turn] No problem! Once you put that in, it should show your desktop in a new window. If that's not working, let me know what error you see and we'll troubleshoot from there. ▷ yields, end



Transcript of test_00169

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–14.2 **S1** [turn] Alright fam, so here's my master plan for Sunday! First, we get everyone up by eight, even you, Dad—no sleeping in!—because we need to pack lunches and get on the road by nine sharp, or else we'll hit traffic on the way to grandma's. ▷ yields to S3

4.6–6.5 **S2** [failed claim] Yeah, hey, I resent that!

11.2–13.9 **S4** [laughter] Haha, good luck getting him out of bed that early!

14.1–16.6 **S3** [turn] Eight isn't bad. What's the plan if we run late? ▷ yields to S1

16.6–23.7 **S1** [turn] If we run late, we, uh, skip Starbucks and go straight to grandma's. No detours—seriously! I know, I know, that's tragic.

20.2–20.8 **S2** [failed claim] Wait—

23.8–26.5 **S1** [turn] —but it's the only way we make it for brunch, honestly. ▷ yields to S2

26.4–32.3 **S2** [turn] I was just gonna say, can't we order Starbucks on the app and do pickup? That way we don't have to wait in line if we're late. ▷ yields to S1

28.0–28.6 **S4** [backchannel] mm-hm

30.4–31.2 **S3** [backchannel] Good idea.

32.7–41.5 **S1** [turn] Fine, you know what, that's actually genius. Okay, so Starbucks app for Dad. Then, once we're at grandma's, please, can someone help me unload all the stuff? ▷ yields to S4

41.4–45.2 **S4** [turn] I can! What exactly are we bringing again? Just the pies and the games? ▷ yields to S3

45.1–48.9 **S3** [turn] And the folding chairs. Don't forget, last time there weren't enough for everyone. ▷ yields to S2

48.9–53.5 **S2** [turn] Uh, do we need ice packs for the pies? They melted last year and made that mess in the trunk. ▷ yields to S1

53.7–59.1 **S1** [turn] Ooh, thank you—yes! I'll freeze a couple tomorrow night. See, I knew it was smart to consult the group! ▷ cut off

56.1–57.1 **S2** [backchannel] Sounds good.

57.6–58.2 **S3** [backchannel] Perfect.

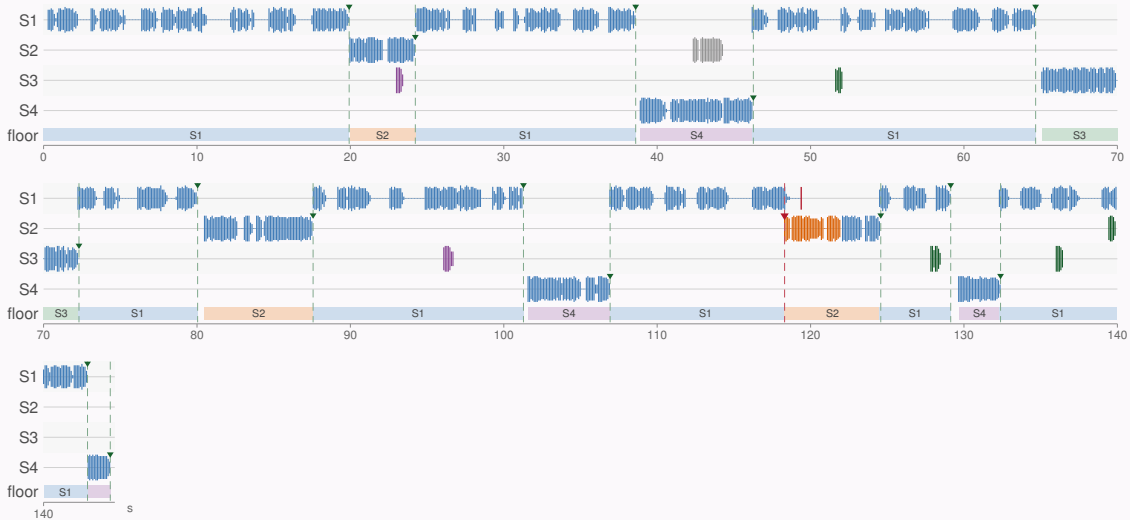
57.9–58.6 **S4** [backchannel] Love it!

58.7–62.8 **S4** [interruption] Oh! Can we carpool back with Aunt Lisa, or should we all ride together again? ▷ takes floor from S1 ▷ holds, pauses

63.8–68.3 **S4** [turn] She said she might have extra space, and if you want, I can confirm with her tonight. ▷ yields to S1

68.3–75.1 **S1** [turn] Oh that'd be awesome, yeah, could you? Let me know what she says, if it's easier, we can definitely split up for the ride home. ▷ yields, end

test split · tone: curious and inquisitive · noise: living room tv distant, 31.6 dB SNR · voices: S1 F, S2 F, S3 M, S4 M
EOT 15 yields/0 holds · INT 1 interruptions/6 negatives (4 backchannels, 2 laughter) · 1 failed claims (excluded)



Transcript of test_00176

start-end (s) speaker [ground-truth label] text > outcome

0.3–19.9 **S1** [turn] Alright, so if everyone could, uh, grab one of those clipboards, I'll walk you through how we sign folks up for shifts at the booth. It's not rocket science, but you do have to write legibly, which seems to be a big ask around here. > yields to S2

20.0–24.2 **S2** [turn] Ooh, do we get to pick the color of our pens, or is that too advanced for today? > yields to S1

23.0–23.5 **S3** [laughter] haha

24.3–38.6 **S1** [turn] If I had any control over pens, I'd be rich. Just, uh, write your names at the top right corner—that's for tracking in case anyone tries to escape. > yields to S4

38.9–44.3 **S4** [turn] So, um, after we write our names—do we, like, choose our own time slot, or is there a set schedule already?

42.4–44.3 **S2** [failed claim] Yeah, what if I only want the donut hour?

44.4–46.3 **S4** [turn] —uh, I mean, just asking for a friend. > yields to S1

46.2–64.7 **S1** [turn] Believe it or not, the donut hour is hotly contested. But yeah, you pick your slot—just write your name next to the time you want. If someone's already there, you can double up or find another slot. > yields to S3

51.6–52.1 **S3** [backchannel] uh-huh

65.1–72.3 **S3** [turn] So, um, if two people want the same slot, is there any kind of, uh, priority? Or is it just first come, first served? > yields to S1

72.2–80.0 **S1** [turn] Mostly first come, first served. Unless you bribe me with coffee, then all bets are off. > yields to S2

80.5–87.6 **S2** [turn] Wait, what if we need to swap after signing up? Like if, uh, I realize I'm double-booked for yoga and origami somehow? > yields to S1

87.6–101.3 **S1** [turn] You can swap, but you need to actually tell someone. Preferably me. Preferably not at three in the morning via frantic texts, but, you know, life happens. > yields to S4

96.1–96.7 **S3** [laughter] haha right

101.6–106.9 **S4** [turn] Um, is there anything we need to do during the shift, like specific tasks? Or is it just... hang out at the booth? > yields to S1

106.9–119.4 **S1** [turn] Great question—so mainly you greet people, hand out flyers, and collect names for the newsletter. If someone asks a wild question, just nod thoughtfully and come get me. > cut off

118.3–122.0 **S2** [interruption] Oh! What if someone asks about volunteering, like, for the next event? > takes floor from S1

122.1–124.6 **S2** [turn] Do we recruit them right then, or... point them at you? > yields to S1

124.5–129.1 **S1** [turn] Recruit them! And then point them at me. Double victory. > yields to S4

127.8–128.5 **S3** [backchannel] makes sense

129.7–132.4 **S4** [turn] uh, thanks—this actually seems, um, doable. Even for me. > yields to S1

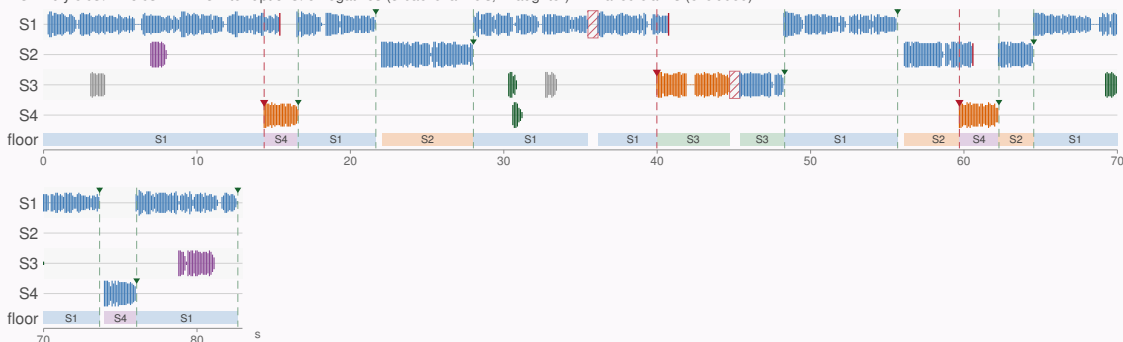
132.3–142.9 **S1** [turn] Trust me, if you survive this, you can coordinate anything. Questions before I start assigning people to balloon duty? > yields to S4

136.0–136.5 **S3** [backchannel] um, I'm good

139.4–139.9 **S2** [backchannel] nope

142.9–144.3 **S4** [turn] No, I think I'm okay. > yields, end

test split · tone: enthusiastic and energetic · noise: dorm room, 25.3 dB SNR; paper shuffle, laughter offmic · voices: S1 M, S2 M, S3 F, S4 F
EOT 10 yields/2 holds · INT 3 interruptions/5 negatives (3 backchannels, 2 laughter) · 2 failed claims (excluded)



Transcript of test_00177

start-end (s) speaker [ground-truth label] text > outcome

0.3–8.8 **S1** [turn] Alright, so—uh, welcome, you guys! Today I'm gonna walk you through how to schedule a medical appointment in our system. It's honestly way easier than people think, promise.

3.1–4.1 **S3** [failed claim] Love to hear that!

7.0–8.1 **S2** [laughter] Haha, yeah, let's hope.

8.9–15.4 **S1** [turn] So first thing, you log into the portal with your work credentials. You'll see a dashboard, and there's this big blue button, says 'Request Appointment.' You can't miss it. > cut off

14.4–16.6 **S4** [interruption] And that works on mobile? Or desktop only? > takes floor from S1 > yields to S1

16.5–21.7 **S1** [turn] Great question! Both, actually. Website works on your phone or on a computer—super flexible. > yields to S2

22.0–28.0 **S2** [turn] Okay, so after you click that button, what do you have to fill out? Like, does it ask a million questions or is it quick? > yields to S1

28.0–35.5 **S1** [turn] It's quick! They just want basics—doctor or department, your preferred dates and times, and a super short reason for your visit, like just a couple words. > holds, pauses

30.3–30.8 **S3** [backchannel] Nice!

30.6–31.2 **S4** [backchannel] Right.

32.7–33.5 **S3** [failed claim] Wait, so if—

36.1–40.7 **S1** [turn] —and then after you submit, you'll get a confirmation email pretty quick—sometimes within minutes, or up to an hour if it's busy. > cut off

40.0–44.7 **S3** [interruption] So what if you need to reschedule after that—do you use the site again or call the office? > takes floor from S1 > holds, pauses

45.4–48.3 **S3** [turn] Like, is there a giant red reschedule button or something? > yields to S1

48.3–55.7 **S1** [turn] Haha, not red, but yeah! There's a 'Manage Appointments' tab—super easy to find. Just choose your appointment there, click reschedule, and pick a new time. > yields to S2

56.1–60.6 **S2** [turn] Alright, that's honestly way smoother than I thought. And, uh, what about just straight up canceling if you can't make it last minute? > cut off

59.7–62.3 **S4** [interruption] Yeah, do we have to call if it's close to the appointment? > takes floor from S2 > yields to S2

62.2–64.5 **S2** [turn] Or can you just click cancel even an hour before? > yields to S1

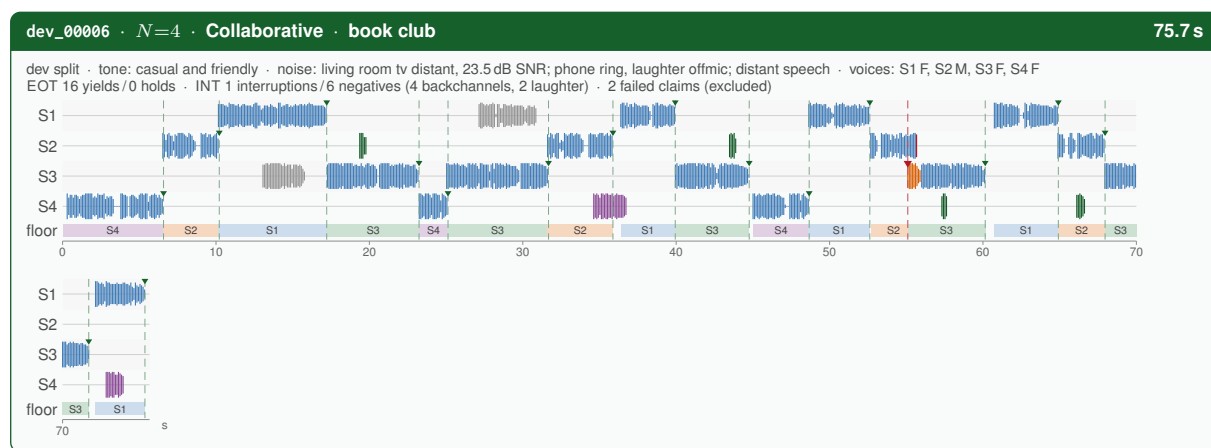
64.5–73.7 **S1** [turn] Good one—so most of the time, you can just cancel online, super simple. If it's under 24 hours though, it'll say you need to call the office, but otherwise, just one click. > yields to S4

69.2–70.0 **S3** [backchannel] Works for me!

74.0–76.1 **S4** [turn] Alright. Thanks for clarifying everything. > yields to S1

76.0–82.7 **S1** [turn] Of course! And honestly, ping me any time if you get stuck or just wanna vent about weird appointment messages—I'm always around! > yields, end

78.8–81.2 **S3** [laughter] Haha, hope you're ready for all my questions then!



Transcript of dev_00006

start-end (s) speaker [ground-truth label] text > outcome

0.3–6.6 **S4** [turn] Okay, um, so did anyone actually finish the book this time? I mean, be honest, I barely got through the last third. > yields to S2

6.5–10.2 **S2** [turn] I finished, but not out of enjoyment. It felt like a chore. > yields to S1

10.2–14.0 **S1** [turn] Seriously? I mean, was it just me or was the plot, like, weirdly circular?

13.1–15.8 **S3** [failed claim] You think so? I thought the structure was part of the point—

14.1–17.2 **S1** [turn] —like, she kept repeating herself but not adding anything new. > yields to S3

17.3–23.2 **S3** [turn] Right, but I think it was showing how she was stuck, emotionally? That repetition, it mirrored her mental state. > yields to S4

19.4–19.9 **S2** [backchannel] mm-hm

23.3–25.1 **S4** [turn] Oh, I hadn't thought of it like that. > yields to S3

25.0–29.8 **S3** [turn] Yeah, and the flashbacks—those weren't just filler, right? Like, they were how she processed the guilt.

27.2–30.9 **S1** [failed claim] But wait, if it's all about guilt, why does she push everyone away? Wouldn't she want—

29.9–31.7 **S3** [turn] —I think that's the defense mechanism. > yields to S2

31.6–35.9 **S2** [turn] She can't connect, so she does the opposite. Like a... reverse magnet. > yields to S1

34.6–36.8 **S4** [laughter] Haha, a reverse magnet, I love that.

36.4–39.9 **S1** [turn] But doesn't that get exhausting to read though? I felt, I don't know, shut out. > yields to S3

40.0–44.8 **S3** [turn] Absolutely, but maybe that's why it hits so hard? You're supposed to feel her isolation. > yields to S4

43.5–44.0 **S2** [backchannel] yeah

45.0–48.7 **S4** [turn] I, um, just wanted more out of the ending. It felt abrupt, right? > yields to S1

48.6–52.6 **S1** [turn] Yes! Oh my god, thank you. I got to the last page and was like, wait, that's it? > yields to S2

52.7–55.7 **S2** [turn] Same, I literally flipped the page looking for an epilogue. > cut off

55.1–55.9 **S3** [interruption] No, but I think— > takes floor from S2

56.0–60.1 **S3** [turn] —maybe that's intentional, you know? Like, we don't get closure because she doesn't. > yields to S1

57.3–57.7 **S4** [backchannel] ohh

60.7–64.9 **S1** [turn] But isn't that kind of a cop-out, though? Like, sometimes I just want the author to commit. > yields to S2

64.9–68.0 **S2** [turn] I mean, life doesn't usually end tidy either. > yields to S3

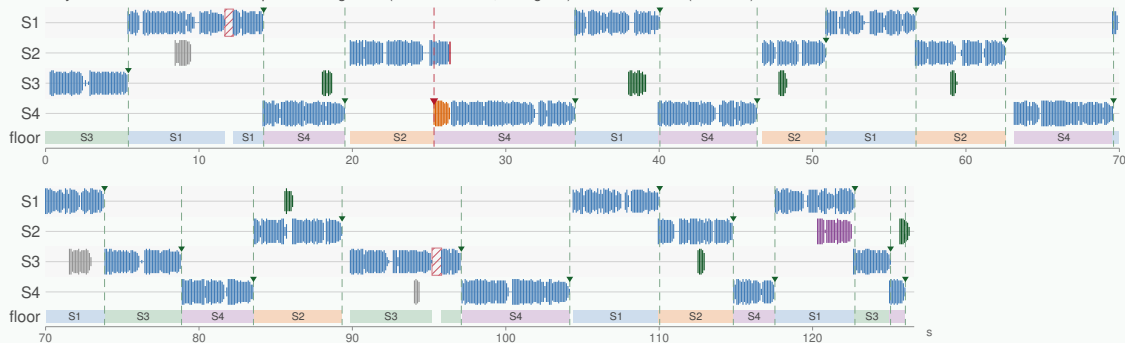
66.1–66.7 **S4** [backchannel] mm-hm

67.9–71.7 **S3** [turn] Right, but it does beg the question—would you guys read something else by her? > yields to S1

72.1–75.4 **S1** [turn] Honestly? Only if you convinced me it'd be less frustrating. > yields, end

72.8–74.0 **S4** [laughter] Haha, that's fair.

dev split · tone: curious and inquisitive · noise: school hallway, 36.0dB SNR; door knock · voices: S1 F, S2 M, S3 M, S4 F
EOT 22 yields / 2 holds · INT 1 interruptions / 8 negatives (7 backchannels, 1 laughter) · 3 failed claims (excluded)

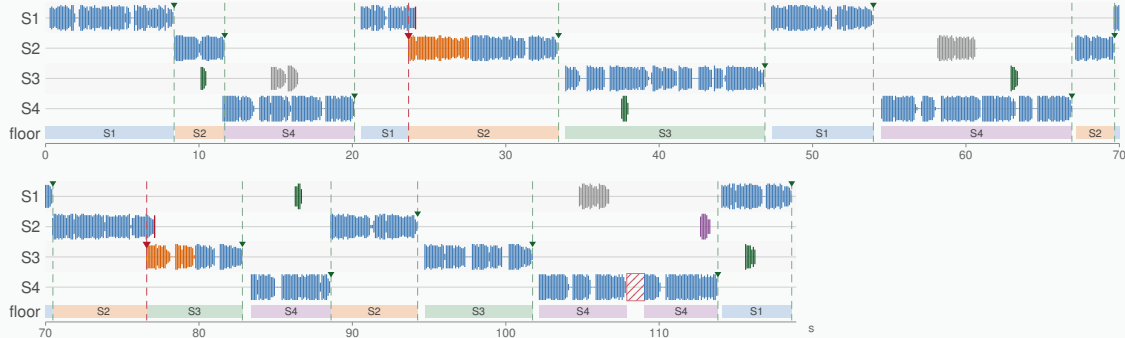


Transcript of dev_00021

start—end (s) speaker [ground-truth label] text > outcome

0.3–5.4 **S3** [turn] Alright, so we're all here— Are we starting with the kitchen? That seemed to be the, uh, biggest pain point last time. > yields to S1
5.4–11.7 **S1** [turn] I think so! And I love the idea of knocking down that wall by the breakfast nook, just to open everything up more. > holds, pauses
8.5–9.6 **S2** [failed claim] But structurally—
12.2–14.2 **S1** [turn] —I know it's possibly load-bearing, right? > yields to S4
14.2–19.5 **S4** [turn] Yes, exactly. We should get an engineer to verify before we actually plan for demolition. > yields to S2
18.0–18.7 **S3** [backchannel] Makes sense.
19.9–26.4 **S2** [turn] If we do end up removing that wall, I want to talk about the peninsula idea. What if instead we did an island? I measured— > cut off
25.3–26.4 **S4** [interruption] Wait, I was going to say— > takes floor from S2
26.5–34.5 **S4** [turn] The way the plumbing is currently routed, a full island would need us to extend lines under the floor. That's, um, a big cost jump. > yields to S1
34.5–40.0 **S1** [turn] Right, but like, isn't that also true if we try to move the sink at all? Or is the peninsula less tricky? > yields to S4
38.0–39.2 **S3** [backchannel] Yeah, good question.
39.9–46.4 **S4** [turn] Peninsula's a bit easier since it's closer to the existing lines. Still, you're right—it's a project either way. > yields to S2
46.7–50.9 **S2** [turn] So if cost is the deciding factor, peninsula it is for now? > yields to S1
47.8–48.3 **S3** [backchannel] Uh-huh.
50.9–56.7 **S1** [turn] Yeah, bummer, but—wait, what about resale value? Is an island, like, really that much better for future buyers? > yields to S2
56.7–62.6 **S2** [turn] Depends a lot on how much seating you can fit. If it's just a little one, peninsula might win. > yields to S4
59.0–59.4 **S3** [backchannel] Right.
63.1–69.6 **S4** [turn] It's also about the flow. With an island you might get bottlenecks around the fridge if we aren't careful with placement. > yields to S1
69.6–72.3 **S1** [turn] Oh! And I wanted to, um, ask about the lighting—
71.6–73.0 **S3** [failed claim] Actually, before we jump—
72.3–73.8 **S1** [turn] —Sorry, yeah, go ahead. > yields to S3
73.9–78.9 **S3** [turn] Just, uh, we didn't really talk about flooring yet. If we're gutting so much, do we go hardwood or tile? > yields to S4
78.9–83.6 **S4** [turn] Hardwood scratches more but it feels so much warmer. In a kitchen though, I lean tile. > yields to S2
83.6–89.3 **S2** [turn] Yes, plus with the open concept, matching the rest of the house might matter for continuity. > yields to S3
85.6–86.2 **S1** [backchannel] Mhm.
89.9–95.2 **S3** [turn] But like, what about that cork option you showed last week? Is it... I dunno, is it durable enough for splashes? > holds, pauses
94.0–94.4 **S4** [failed claim] Well—
95.8–97.1 **S3** [turn] —'Cause it looks awesome. > yields to S4
97.2–104.2 **S4** [turn] It does, but you'd definitely want to seal it really well. Otherwise water damage is a risk, especially around the sink. > yields to S1
104.4–110.0 **S1** [turn] Okay, but if we did that and stuck with tile by the sink for, like, a small border, would that be weird? > yields to S2
109.9–114.8 **S2** [turn] Not necessarily. Actually, that could look very custom if we design it right. > yields to S4
112.5–113.0 **S3** [backchannel] Nice.
114.9–117.5 **S4** [turn] Okay, I'll sketch up some options there. Now, lighting— > yields to S1
117.6–121.4 **S1** [turn] Yes! Pendant lights over the peninsula! Can we do, like, three glass ones?
120.3–122.6 **S2** [laughter] Haha, you love your glass lighting.
121.5–122.8 **S1** [turn] It's so pretty though! > yields to S3
122.7–125.1 **S3** [turn] Let's just— can we try a dimmer this time? Please? > yields to S4
125.0–126.0 **S4** [turn] Absolutely, I'll note it. > yields, end
125.7–126.3 **S2** [backchannel] Great.

test split · tone: curious and inquisitive · noise: quiet apartment, 22.5 dB SNR; chairs scraping · voices: S1 M, S2 M, S3 F, S4 M
EOT 15 yields / 1 holds · INT 2 interruptions / 6 negatives (5 backchannels, 1 laughter) · 3 failed claims (excluded)

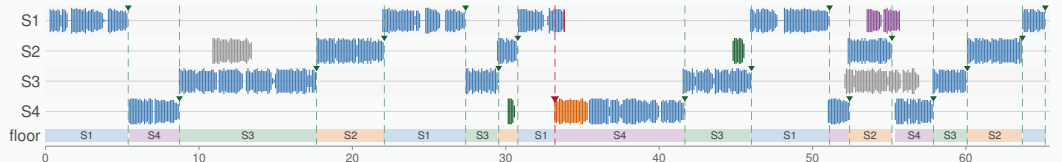


Transcript of test_00030

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.4 **S1** [turn] Um, so, I guess I'll start? Uh, has anyone read The Midnight Library by Matt Haig? I finished it last week and—it's kind of stuck with me. ▷ yields to S2
8.4–11.7 **S2** [turn] Oh yeah, I've seen that one around. Like, it's got the blue cover, right? ▷ yields to S4
10.1–10.5 **S3** [backchannel] mm-hm
11.6–16.0 **S4** [turn] Yeah, I've heard the concept's really interesting. It's about, uh, parallel lives, right?
14.7–16.5 **S3** [failed claim] But isn't it kind of, I don't know—
16.1–20.1 **S4** [turn] —a little on-the-nose with the metaphors? I haven't read it, but that's what I've heard. ▷ yields to S1
20.6–24.1 **S1** [turn] Uh, maybe, but I thought it, like, actually handled regret and choice in a pretty honest way. ▷ cut off
23.7–27.7 **S2** [interruption] So wait, does it remind anyone else of that movie Sliding Doors or is that— ▷ takes floor from S1
27.7–33.4 **S2** [turn] You know, where Gwyneth Paltrow's life splits based on catching a train? It kinda sounds like the same vibe. ▷ yields to S3
33.9–41.2 **S3** [turn] Oh, totally! I was gonna say that, like, the whole alternate timeline thing has been done, but I guess if the writing's good—
37.6–38.0 **S4** [backchannel] yeah
41.3–46.9 **S3** [turn] —it still works. Speaker_1, did you, like, feel it was original or more like a comfort read? ▷ yields to S1
47.3–54.0 **S1** [turn] Um, I think more comfort-y, but the library itself is kind of cool? I just liked the little details, honestly. ▷ yields to S4
54.5–61.0 **S4** [turn] So if you liked that, have you read Life After Life? Kate Atkinson? The idea's kinda similar, but maybe more literary.
58.2–60.7 **S2** [failed claim] Is that the one where she, like, keeps being reborn, or—
61.1–66.9 **S4** [turn] —yeah, it resets every time she dies. It's heavier, though. I think you'd like it if you liked Midnight Library. ▷ yields to S2
62.9–63.5 **S3** [backchannel] oh wow
67.2–69.7 **S2** [turn] Alright, if we're swapping, can I toss out a movie instead? ▷ yields to S1
69.7–70.5 **S1** [turn] Yeah, go for it. ▷ yields to S2
70.5–77.1 **S2** [turn] So, Coherence? It's, like, this super low-budget sci-fi where friends at a dinner party get sucked into alternate realities. It's wild. ▷ cut off
76.6–79.7 **S3** [interruption] I've never even heard of that one. Is it confusing or— ▷ takes floor from S2
79.8–82.8 **S3** [turn] Does it actually make sense, or is it just, like, chaos? ▷ yields to S4
83.4–88.6 **S4** [turn] I think it's pretty clever, actually. It's got that puzzle-box thing, but it ties up a lot by the end. ▷ yields to S2
86.3–86.7 **S1** [backchannel] uh-huh
88.6–94.3 **S2** [turn] Yeah, you really have to pay attention, but it's worth it. And it's just, like, kind of unsettling in a fun way. ▷ yields to S3
94.7–101.7 **S3** [turn] I feel like all these jump between books and movies. Anybody got something totally different? Like a documentary, maybe? ▷ yields to S4
102.2–107.9 **S4** [turn] Has anyone seen Won't You Be My Neighbor? The Mister Rogers one? It's a whole different mood, but it's so good. ▷ holds, pauses
104.8–106.8 **S1** [failed claim] I cried at that, honestly. Like, it was—
109.0–113.8 **S4** [turn] —yeah, it really gets you. It's such a palate cleanser after all the, uh, existential stuff. ▷ yields to S1
112.7–113.4 **S2** [laughter] haha right
114.1–118.6 **S1** [turn] I think that's my favorite type of recommendation—the ones you don't expect, you know? ▷ yields, end
115.6–116.3 **S3** [backchannel] exactly

test split · tone: casual and friendly · noise: quiet apartment, 21.5 dB SNR; phone ring · voices: S1 F, S2 F, S3 F, S4 M
EOT 16 yields / 0 holds · INT 1 interruptions / 3 negatives (2 backchannels, 1 laughter) · 2 failed claims (excluded)

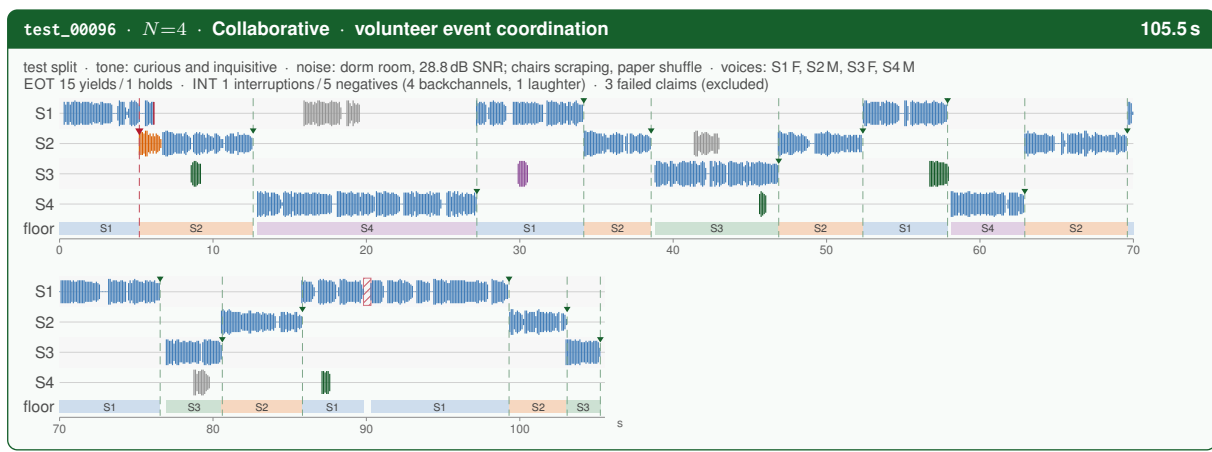


Transcript of test_00072

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–5.4 **S1** [turn] Okay, so, uh, are we all trying to get our flu shots this week or what's the game plan here? ▷ yields to S4
5.4–8.7 **S4** [turn] I mean, if we can get the same slot, that'd be super easy honestly. ▷ yields to S3
8.7–12.9 **S3** [turn] Yeah, totally, I can coordinate if you want. Just let me know your windows?
10.9–13.5 **S2** [failed claim] Wait, are we doing this together or are people just—
13.1–17.7 **S3** [turn] —trying to sync up two at a time? 'Cuz I think the clinic won't do like, you know, a big group. ▷ yields to S2
17.7–22.1 **S2** [turn] No yeah, I just meant, like, I can come over lunch break but it can't be late afternoon for me. ▷ yields to S1
22.0–27.4 **S1** [turn] Haha, I can do basically any time before four, except Wednesdays—Tuesdays are chaos. ▷ yields to S3
27.4–29.5 **S3** [turn] Right, so morning maybe works better? ▷ yields to S2
29.5–30.8 **S2** [turn] Yeah, that's fine for me. ▷ yields to S1
30.2–30.6 **S4** [backchannel] Mm-hmm
30.8–33.8 **S1** [turn] Can we check online if there's, like, a slot for all of us at, say, eleven? ▷ cut off

33.2–35.4 **S4** [interruption] I can pull up the portal right now while you guys— **▷ takes floor from S1**
 35.5–41.7 **S4** [turn] Yeah, okay, so it looks like there are like, three back-to-back at 11:10, 11:20, and 11:30. **▷ yields to S3**
 41.6–46.0 **S3** [turn] Oh nice! That's basically perfect. I can do 11:10 if nobody minds. **▷ yields to S1**
 44.8–45.6 **S2** [backchannel] Go for it.
 46.0–51.1 **S1** [turn] Yeah, I'll grab eleven thirty! I need the extra twenty minutes for coffee recovery anyway. **▷ yields to S4**
 51.0–52.4 **S4** [turn] I'll take eleven twenty, easy. **▷ yields to S2**
 52.1–57.0 **S3** [failed claim] So, wait, does someone want to call to make sure they can do three at once or should we just—
 52.3–55.2 **S2** [turn] —oh, I can call. I'm good at being annoying with the front desk. **▷ yields to S4**
 53.6–55.8 **S1** [laughter] Haha, please do! They love you at this point.
 55.4–57.9 **S4** [turn] Seriously, you get things done faster than anyone. **▷ yields to S3**
 57.9–60.1 **S3** [turn] Thanks, Bri. Just let us know what they say? **▷ yields to S2**
 60.1–63.7 **S2** [turn] Will do. If they say no, we'll just stagger it and hang in the waiting room and gossip. **▷ yields to S1**
 63.7–65.2 **S1** [turn] Perfect, I'll bring the snacks. **▷ yields, end**

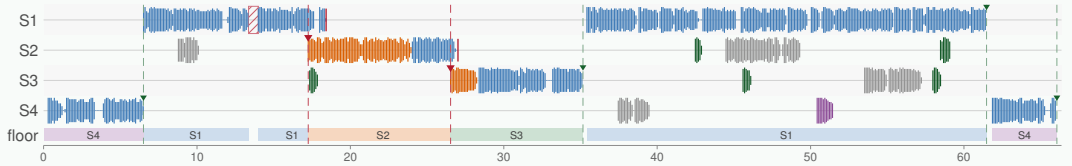


Transcript of test_00096

start–end (s) speaker [ground-truth label] text **▷ outcome**

0.3–6.1 **S1** [turn] Okay, so I know we said we wanted a pancake breakfast again, but what if—hear me out—what if we tried waffles this year? **▷ cut off**
 5.2–6.7 **S2** [interruption] Oh, waffles could be— **▷ takes floor from S1**
 6.7–12.6 **S2** [turn] Like, that's honestly way more exciting! They have those deep holes for toppings. We could do, like, a topping bar maybe? **▷ yields to S4**
 8.6–9.2 **S3** [backchannel] yeah
 12.9–20.6 **S4** [turn] I, uh, I like that idea, but have we checked if the community center even has waffle makers? I think they only had those giant griddles last year.
 15.9–19.6 **S1** [failed claim] Oh! Good point, I didn't even think about equipment. Wait, what if we—
 20.7–27.2 **S4** [turn] —I mean we could always borrow some? Or maybe ask volunteers if they have one to bring. But that could get tricky with cords and stuff. **▷ yields to S1**
 27.2–34.2 **S1** [turn] Haha, yeah, what if someone brings like, a mini Hello Kitty waffle iron and we end up with cat-shaped waffles everywhere? **▷ yields to S2**
 29.9–30.6 **S3** [laughter] haha
 34.2–38.6 **S2** [turn] Actually, that sounds super cute, people would love that! But, um, logistics-wise... **▷ yields to S3**
 38.8–43.5 **S3** [turn] So maybe we should do pancakes and just add more fun toppings instead? Like fruit, sprinkles—
 41.4–43.0 **S2** [failed claim] Or whipped cream mountains! Oh—
 43.6–46.9 **S3** [turn] —yeah, and maybe some, uh, gluten-free batter for people who need it? **▷ yields to S2**
 45.6–46.1 **S4** [backchannel] mm-hm
 46.9–52.4 **S2** [turn] Right! We could set up, like, two stations? One with the normal batter and one gluten-free so there's no mix-up? **▷ yields to S1**
 52.4–57.9 **S1** [turn] Yes! And maybe put up cute signs or something? Like, little pancake characters showing which is which? **▷ yields to S4**
 56.7–58.0 **S3** [backchannel] aww, that's adorable
 58.1–62.9 **S4** [turn] Hang on, uh, are we asking volunteers to bring all the toppings too, or should we budget for that? **▷ yields to S2**
 62.9–69.6 **S2** [turn] I think, hm, maybe we can do a sign-up sheet? Like, people who want to donate toppings can write it in and then we'll know what to buy ourselves. **▷ yields to S1**
 69.7–76.6 **S1** [turn] Ooh, can we make it one of those online forms this time? Last year my spreadsheet just devolved into colorful chaos. **▷ yields to S3**
 77.0–79.2 **S3** [turn] I can help set up a Google Form, if you'd like?
 78.8–79.8 **S4** [failed claim] Thank you, that's—
 79.3–80.6 **S3** [turn] —no problem at all! **▷ yields to S2**
 80.6–85.8 **S2** [turn] Right, and someone should, um, be in charge of collecting allergy info? So no peanut disasters. **▷ yields to S1**
 85.8–89.8 **S1** [turn] That should be me, since I already get, like, five allergy emails anyway. **▷ holds, pauses**
 87.1–87.7 **S4** [backchannel] okay
 90.3–99.3 **S1** [turn] So, to sum up, pancakes with wild toppings, two stations, online sign-ups, and no Hello Kitty appliances unless they're, like, regulation size? **▷ yields to S2**
 99.3–103.1 **S2** [turn] Perfect summary. Uh, can we make the flyer look like a pancake stack too? **▷ yields to S3**
 103.0–105.2 **S3** [turn] I love it, I'll start doodling something later! **▷ yields, end**

test split · tone: enthusiastic and energetic · noise: living room tv distant, 27.6 dB SNR; laughter offmic · voices: S1 F, S2 M, S3 M, S4 M
EOT 4 yields/1 holds · INT 2 interruptions/6 negatives (5 backchannels, 1 laughter) · 4 failed claims (excluded)

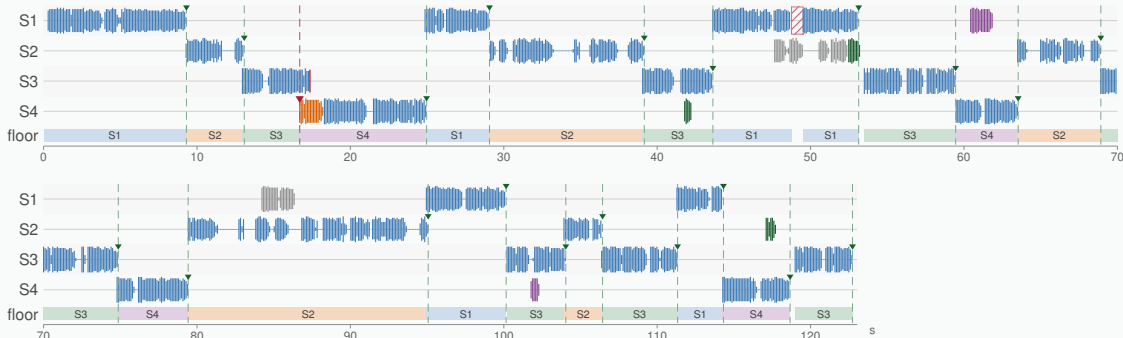


Transcript of test_00134

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.5 **S4** [turn] Alright everyone, let's give this pitch run-through a shot. Whenever you're ready, go ahead and kick us off, Maya. ▷ yields to S1
6.5–13.4 **S1** [turn] Thanks. Okay—imagine a world where finding a reliable dog walker is as easy as ordering coffee. That's what we're building with Waggle. ▷ holds, pauses
8.8–10.1 **S2** [failed claim] Uh, just to clarify, are we—
14.0–18.4 **S1** [turn] —creating a seamless way for pet owners to connect with background-checked walkers, in real time, through an app that's genuinely user-friendly. ▷ cut off
17.3–24.0 **S2** [interruption] Okay, sorry, so are we focusing only on urban areas, or is this something we see rolling out to suburbs too? ▷ takes floor from S1
17.3–17.9 **S3** [backchannel] Right.
24.0–27.0 **S2** [turn] Because, um, the demand profiles are really different, and I think— ▷ cut off
26.5–28.2 **S3** [interruption] Wait. I think there's something here— ▷ takes floor from S2
28.4–35.2 **S3** [turn] Like, should our pitch highlight how we're uniquely tailoring for, um, high-density cities first? Because that gives us a clearer go-to-market. ▷ yields to S1
35.4–40.1 **S1** [turn] That makes sense—yeah, we can position it as urban-focused at launch, then expand as we learn.
37.5–39.5 **S4** [failed claim] Nice adjustment. That's punchier.
40.2–48.4 **S1** [turn] So, our core value prop is trust and convenience. We pre-vet every walker, and scheduling's instant. That's something no other app is, um, really nailing right now.
42.5–42.9 **S2** [backchannel] Yeah.
44.5–49.3 **S2** [failed claim] Can I just—so would it help if we had a real user story up front? Like, an anecdote?
45.6–46.1 **S3** [backchannel] Mhm.
48.5–55.4 **S1** [turn] —Yeah, that's a great idea! Like, um, 'Sarah's a nurse with weird hours, she can't find a walker she trusts on short notice—until Waggle.'
50.4–51.5 **S4** [laughter] Haha, perfect.
53.5–57.3 **S3** [failed claim] What about the competition slide? Are we, uh, calling out Rover by name? Or—
55.5–61.5 **S1** [turn] —I think we have to, right? But it's more like, 'Waggle's faster, safer, and has better coverage in cities,' not just another app. ▷ yields to S4
58.0–58.5 **S3** [backchannel] Right.
58.5–59.2 **S2** [backchannel] Exactly.
61.9–66.1 **S4** [turn] Alright, let's try running that from the top, with the anecdote first. Ready? ▷ yields, end

test split · tone: curious and inquisitive · noise: quiet office, 32.4 dB SNR; laughter offmic, door knock · voices: S1 M, S2 F, S3 M, S4 M
EOT 20 yields/1 holds · INT 1 interruptions/5 negatives (3 backchannels, 2 laughter) · 2 failed claims (excluded)

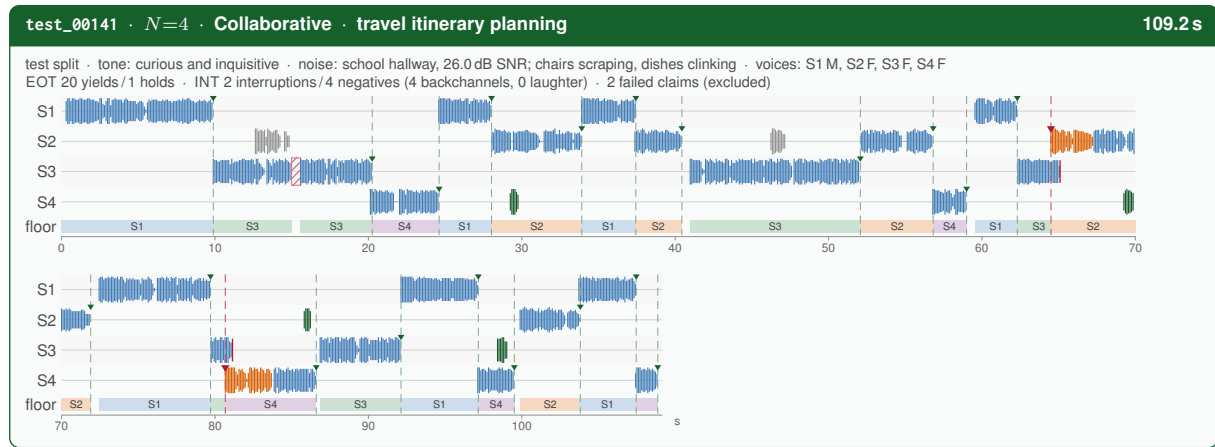


Transcript of test_00137

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.3 **S1** [turn] So, uh, I guess I'll start— this time I actually finished the book. Not to brag. I kinda loved how weird the timeline was, like, with all the jumping back and forth. ▷ yields to S2
9.3–13.1 **S2** [turn] You mean you noticed the timeline was weird? Shocking. ▷ yields to S3
13.0–17.4 **S3** [turn] No, I totally agree, though. The way the author scrambled everything, it made me question— ▷ cut off
16.7–18.2 **S4** [interruption] Wait, before you go on— ▷ takes floor from S3
18.3–25.0 **S4** [turn] Did anyone else get actually confused in chapter five? I had to reread the whole section with the— with the birthday party flashback. ▷ yields to S1
24.9–29.1 **S1** [turn] Yeah, that part threw me. I couldn't tell if the grandmother was alive or a memory. ▷ yields to S2
29.1–39.2 **S2** [turn] Right. And then that random dog shows up out of nowhere. Like, was that magical realism or just a dog? ▷ yields to S3
39.0–43.6 **S3** [turn] I think it was meant to blur the line, honestly. The unreliable narrator thing, maybe? ▷ yields to S1
41.8–42.3 **S4** [backchannel] mm-hm
43.6–48.8 **S1** [turn] But do you think the narrator was unreliable or just, like, confused? There's a difference, I think. ▷ holds, pauses
47.7–52.4 **S2** [failed claim] Oh come on, you can be both. I mean, I felt like they—
49.5–53.1 **S1** [turn] —they were deliberately giving you just enough to be lost, like on purpose. ▷ yields to S3
52.5–53.3 **S2** [backchannel] Exactly.
53.5–59.5 **S3** [turn] I kept waiting for a reveal that tied it all together. Instead, it was like, here's another puzzle piece, good luck. ▷ yields to S4
59.5–63.5 **S4** [turn] Yeah, but didn't you appreciate that? Like, I liked not knowing for sure what was real. ▷ yields to S2
60.4–61.9 **S1** [laughter] Haha, you would. You like chaos.
63.5–68.9 **S2** [turn] Honestly, I was half-convinced the dog was a ghost or something? ▷ yields to S3
68.9–74.9 **S3** [turn] Did anyone notice the colors kept getting mentioned in those scenes? Like, the walls were always changing colors. ▷ yields to S4
74.8–79.4 **S4** [turn] I did—well, sort of. I wasn't sure if it meant something symbolically or if it was just scenery. ▷ yields to S2
79.4–91.3 **S2** [turn] Yeah, that's exactly what I mean. Like, the color blue, specifically, kept showing up whenever the main character remembered the past.

84.2–86.4 **S1** [failed claim] Oh! I totally missed that. Wait, remember when—
 91.5–95.1 **S2** [turn] —the suitcase is described as blue. Uh-huh. $\triangleright$ yields to S1
 95.0–100.2 **S1** [turn] Yeah, and the jacket and the sky in that memory. Maybe the blue's, like, nostalgia or sadness? $\triangleright$ yields to S3
 100.2–104.0 **S3** [turn] Possibly. Or it could be hope, though that feels too optimistic for this book. $\triangleright$ yields to S2
 101.8–102.3 **S4** [laughter] haha
 103.9–106.4 **S2** [turn] That's the bleakest thing you've said all week. $\triangleright$ yields to S3
 106.4–111.3 **S3** [turn] So, um, here's a question—did anyone actually like the ending? Or did it just... stop for you, too? $\triangleright$ yields to S1
 111.3–114.3 **S1** [turn] I mean, it stopped mid-sentence, basically. That was wild. $\triangleright$ yields to S4
 114.3–118.7 **S4** [turn] I liked it. I like when things aren't all wrapped up. Real life doesn't have neat endings. $\triangleright$ yields to S3
 117.1–117.8 **S2** [backchannel] oh wow
 119.0–122.7 **S3** [turn] Fair point. But I am gonna need, like, a palate cleanser book next month. $\triangleright$ yields, end

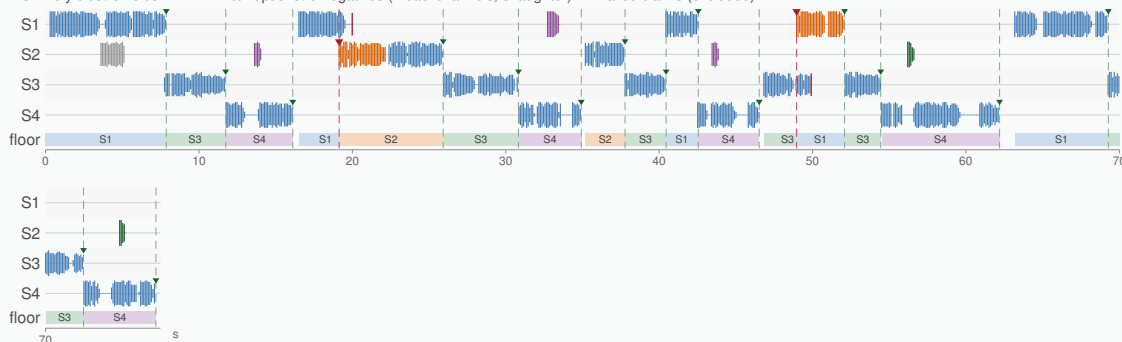


Transcript of test_00141

start–end (s) speaker [ground-truth label] text $\triangleright$ outcome

0.3–9.9 **S1** [turn] So, uh, we're starting in Barcelona, right? Like, we land there at nine in the morning, and then... do we go straight to the hotel or do you wanna drop our bags and wander right away? $\triangleright$ yields to S3
 9.9–15.0 **S3** [turn] I'm honestly good with just leaving our bags at the front desk—most places let you, right? $\triangleright$ holds, pauses
 12.6–14.9 **S2** [failed claim] That's usually fine, yeah. But—
 15.6–20.2 **S3** [turn] —then we can go straight to the market? I feel like food is probably our first stop, right? $\triangleright$ yields to S4
 20.1–24.6 **S4** [turn] That sounds lovely, honestly. The market's not far from where we're staying, I think it's like a— $\triangleright$ yields to S1
 24.6–28.0 **S1** [turn] Yeah, La Boqueria! I looked it up, it's like, a ten minute walk, tops. $\triangleright$ yields to S2
 28.1–33.9 **S2** [turn] Okay but do we want to risk being super tired and hangry? What if we just grab coffee at the airport first— $\triangleright$ yields to S1
 29.3–29.8 **S4** [backchannel] mm-hm
 33.9–37.4 **S1** [turn] No, c'mon, jet lag is part of the adventure. Let's just go straight. $\triangleright$ yields to S2
 37.4–40.4 **S2** [turn] I'm just saying—if one person gets cranky I called it. $\triangleright$ yields to S3
 41.0–47.6 **S3** [turn] Haha, deal. After the market, do we, uh, want to do the Sagrada Familia right away or is that too much for day one?
 46.2–47.2 **S2** [failed claim] No, wait—
 47.8–52.1 **S3** [turn] —it's a lot, but if we do a timed entry for like, midday, maybe that's manageable? $\triangleright$ yields to S2
 52.1–56.8 **S2** [turn] But we should book those tickets ahead, or they'll sell out. I can do that right now actually. $\triangleright$ yields to S4
 56.8–59.0 **S4** [turn] That would take the stress off. Thank you, Mia. $\triangleright$ yields to S1
 59.5–62.3 **S1** [turn] So what about the afternoon? Do we nap, or just fight through? $\triangleright$ yields to S3
 62.3–65.1 **S3** [turn] Nap. I mean, I always say I won't but then I totally— $\triangleright$ cut off
 64.5–67.2 **S2** [interruption] No naps! Power through! Otherwise you're up all night. $\triangleright$ takes floor from S3
 67.3–71.9 **S2** [turn] We can just walk around the Gothic Quarter, maybe stop at a café if someone crashes. $\triangleright$ yields to S1
 69.3–69.9 **S4** [backchannel] yeah
 72.5–79.7 **S1** [turn] Honestly, as long as there's ice cream at some point, I'll survive. Is the beach close by enough for a sunset walk or is that pushing it? $\triangleright$ yields to S3
 79.8–81.1 **S3** [turn] It's maybe twenty minutes? Oh— $\triangleright$ cut off
 80.7–83.7 **S4** [interruption] I think that sounds perfect. Sunset, then tapas, no? $\triangleright$ takes floor from S3
 83.9–86.6 **S4** [turn] We'd probably be ready for a slow dinner by then, right? $\triangleright$ yields to S3
 85.8–86.3 **S2** [backchannel] mm-hm
 86.9–92.1 **S3** [turn] Can we check if there's a flamenco show nearby? Not, like, a huge tourist one, just something chill. $\triangleright$ yields to S1
 92.1–97.2 **S1** [turn] Haha, I'll look. But only if somebody else handles Google Maps, I get us lost every time. $\triangleright$ yields to S4
 97.1–99.5 **S4** [turn] I'll do it. Don't worry, we'll get everywhere together. $\triangleright$ yields to S2
 98.4–99.1 **S3** [backchannel] perfect
 99.9–103.8 **S2** [turn] And someone remind me to book the Sagrada tickets tonight, or I'll forget. $\triangleright$ yields to S1
 103.7–107.5 **S1** [turn] Deal. So... markets, church, beach, food, flamenco, lost—but happily lost. $\triangleright$ yields to S4
 107.4–108.9 **S4** [turn] Sounds just right to me. $\triangleright$ yields, end

test split · tone: casual and friendly · noise: school hallway, 26.2 dB SNR; door knock · voices: S1 M, S2 M, S3 F, S4 F
EOT 16 yields / 0 holds · INT 2 interruptions / 5 negatives (2 backchannels, 3 laughter) · 1 failed claims (excluded)

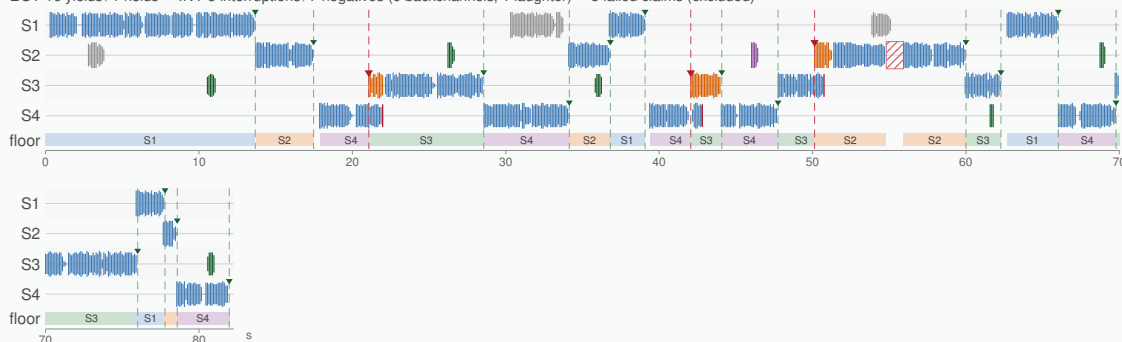


Transcript of test_00167

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–5.7 **S1** [turn] Alright, looks like we need to find a slot for Dr. Lee next week. Should we just pull up the calendar now?
3.6–5.2 **S2** [failed claim] Um, yeah, I can do that, I think, uh—
5.7–7.9 **S1** [turn] —I've got it on the big screen if that's easier. ▷ yields to S3
7.8–11.7 **S3** [turn] Perfect, just tell me the blocks, I'll cross-reference with the referral notes. ▷ yields to S4
11.8–16.1 **S4** [turn] Wild suggestion—what if we don't pick Wednesday for once? ▷ yields to S1
13.6–14.1 **S2** [laughter] haha
16.5–20.0 **S1** [turn] Uh, okay, so Monday Dr. Lee is already booked solid. Tuesday's got a couple of spots after two. ▷ cut off
19.1–22.2 **S2** [interruption] Wait, what about Thursday? Didn't he say he prefers mornings? ▷ takes floor from S1
22.4–25.9 **S2** [turn] I feel like last month he blocked out afternoons for admin stuff. ▷ yields to S3
25.9–30.8 **S3** [turn] Yeah, you're right. Thursday morning is open. There's a 9:30 and an 11. I just checked. ▷ yields to S4
30.9–34.9 **S4** [turn] I call dibs on not calling the patient, by the way. ▷ yields to S2
32.7–33.5 **S1** [laughter] ha, noted
35.2–37.8 **S2** [turn] I can do it, uh, if someone sends me the number again? ▷ yields to S3
37.8–40.5 **S3** [turn] I'll forward it. Should I just text it or email it? ▷ yields to S1
40.4–42.5 **S1** [turn] Text is probably fastest, honestly. ▷ yields to S4
42.5–46.5 **S4** [turn] Honestly, I vote for carrier pigeon, but you do you. ▷ yields to S3
43.5–43.9 **S2** [laughter] haha
46.8–49.9 **S3** [turn] Alright, so Thursday at 9:30? Should we just pencil that in or— ▷ cut off
49.0–52.1 **S1** [interruption] Wait, does the lab need to be notified? Because last time— ▷ takes floor from S3 ▷ yields to S3
52.1–54.4 **S3** [turn] Oh, right, last time the bloods weren't ready. ▷ yields to S4
54.5–62.2 **S4** [turn] Massive plot twist: I actually already cc'd the lab this morning. Gold star for me. ▷ yields to S1
56.2–56.7 **S2** [backchannel] nice
63.2–69.3 **S1** [turn] Okay, looks like we're set then. I'll add it to Dr. Lee's schedule and mark everyone notified. Anything else? ▷ yields to S3
69.3–72.5 **S3** [turn] Nope, I think we actually pulled off a miracle. For once. ▷ yields to S4
72.5–77.2 **S4** [turn] Should we have cake or just quietly bask in this rare moment? ▷ yields, end
74.8–75.2 **S2** [backchannel] mm-hm

test split · tone: enthusiastic and energetic · noise: dorm room, 29.8 dB SNR; phone ring · voices: S1 M, S2 M, S3 F, S4 M
EOT 16 yields / 1 holds · INT 3 interruptions / 7 negatives (6 backchannels, 1 laughter) · 3 failed claims (excluded)

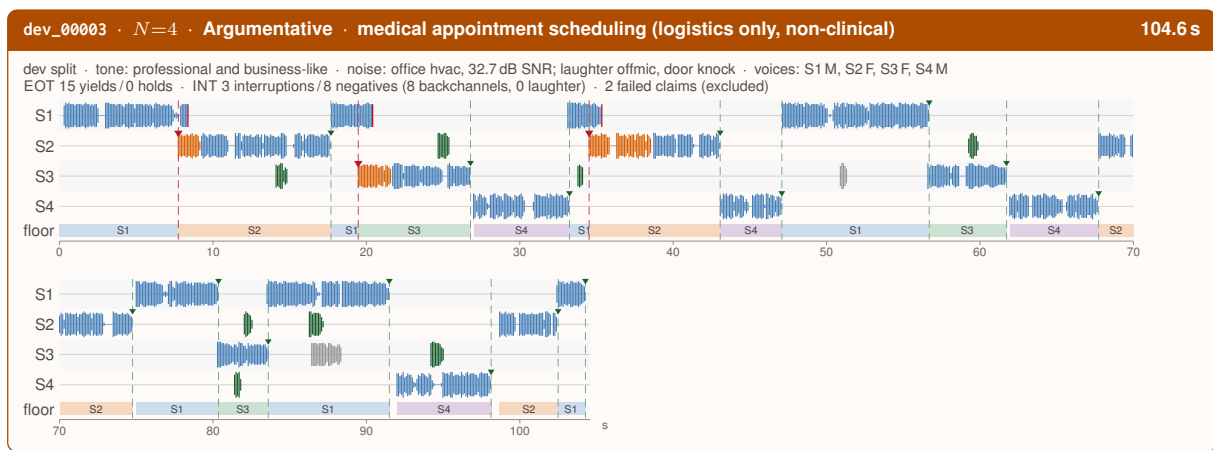


Transcript of test_00170

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.2 **S1** [turn] Um, okay, I'll, uh, start I guess. I thought the part where Maya finally confronts her dad was, um, kinda intense? Like, I didn't see it coming that she'd just—
2.8–3.8 **S2** [failed claim] Oh totally, that scene—
8.2–13.7 **S1** [turn] —she just exploded, you know? I didn't expect that from her after all those chapters of her being super quiet. ▷ yields to S2
10.6–11.1 **S3** [backchannel] mm-hm
13.7–17.5 **S2** [turn] Yeah, like, low-key I was waiting for her to snap. It felt overdue, you know? ▷ yields to S4
17.9–22.0 **S4** [turn] I—uh, I wondered if it was foreshadowed enough? Like, did it feel natural or did it— ▷ cut off
21.1–22.1 **S3** [interruption] I thought so— ▷ takes floor from S4
22.2–28.6 **S3** [turn] She kept bottling everything up, and those journal entries? They felt like little hints she was reaching a breaking point. ▷ yields to S4
26.2–26.7 **S2** [backchannel] right

28.6–31.7 **S4** [turn] Oh, I see, yeah, like the entry after her friend left, that was—
 30.3–33.8 **S1** [failed claim] Wait but did you guys think her dad actually deserved that? I mean—
 31.7–34.1 **S4** [turn] —he seemed, uh, more clueless than cruel to me. > yields to S2
 34.1–36.8 **S2** [turn] No way, he totally gaslit her after the birthday thing. > yields to S1
 35.8–36.3 **S3** [backchannel] yeah
 36.8–39.1 **S1** [turn] Exactly! That was, like, so manipulative. > yields to S4
 39.4–42.8 **S4** [turn] I'm not sure he, uh, fully understood what she needed. He kept saying— > cut off
 42.1–44.1 **S3** [interruption] But that doesn't make it okay, right? > takes floor from S4 > yields to S4
 44.0–47.7 **S4** [turn] —no, no, I agree. I just—maybe he's, um, a product of his own upbringing. > yields to S3
 46.0–46.5 **S2** [laughter] haha
 47.8–50.7 **S3** [turn] I love how we're psychoanalyzing a fictional dad. > cut off
 50.1–51.3 **S2** [interruption] That's the whole point, right? > takes floor from S3
 51.4–54.8 **S2** [turn] Honestly, though, I thought the sister was way more interesting than Maya. > holds, pauses
 53.9–55.1 **S1** [failed claim] Wait, seriously?
 55.9–60.0 **S2** [turn] Yeah, like, she always had an angle. I wanted more chapters from her perspective. > yields to S3
 60.0–62.3 **S3** [turn] Same! Her tension with Maya felt real. > yields to S1
 61.6–61.8 **S4** [backchannel] mm
 62.7–66.0 **S1** [turn] I guess I didn't read her that way, but, um, maybe I missed something. > yields to S4
 66.0–69.8 **S4** [turn] No, I think it was subtle. The author, uh, leaves a lot between the lines. > yields to S3
 68.7–69.1 **S2** [backchannel] yeah
 69.7–76.0 **S3** [turn] Can we talk about the ending? I need to know if I was the only one who, like, flipped back to check if I skipped a page. > yields to S1
 75.9–77.8 **S1** [turn] No, I totally did that too. > yields to S2
 77.7–78.6 **S2** [turn] Same, it was abrupt. > yields to S4
 78.6–82.0 **S4** [turn] I, uh, liked the ambiguity, though. Gave me something to chew on after. > yields, end
 80.6–81.0 **S3** [backchannel] yeah

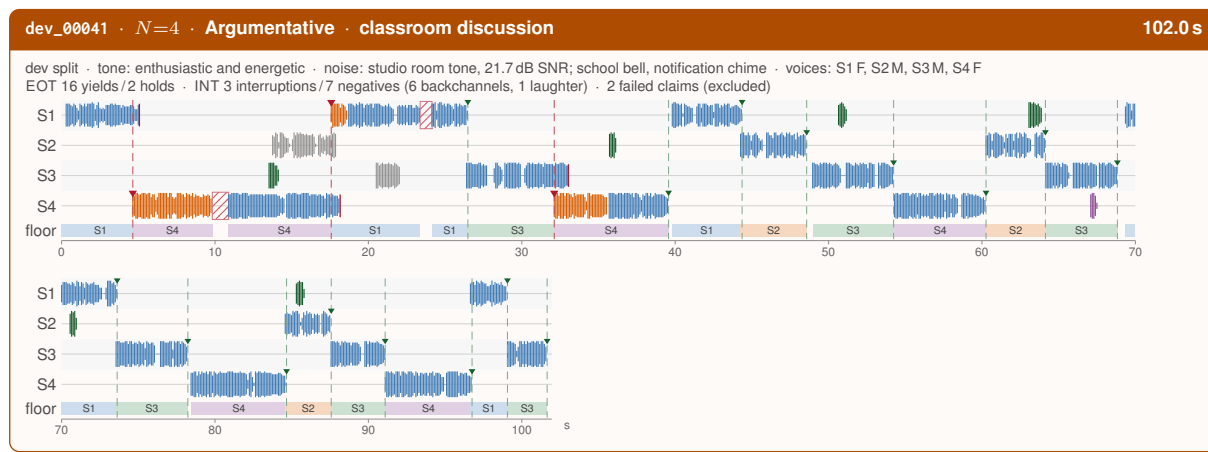


Transcript of dev_00003

start—end (s) speaker [ground-truth label] text > outcome

0.3–8.4 **S1** [turn] So, uh, thanks everyone for jumping on this call. The main thing is we need to figure out how to clear the patient backlog for Dr. Patel's schedule. She's already double-booked most afternoons. > cut off
 7.7–9.2 **S2** [interruption] I'm going to stop you right there— > takes floor from S1
 9.3–17.7 **S2** [turn] The issue isn't just double-booking, it's that pre-authorization times are pushing patients to reschedule, so every fix to the calendar is temporary. > yields to S1
 14.1–14.9 **S3** [backchannel] Exactly
 17.7–20.4 **S1** [turn] But we can't control insurance delays, so we have to work around them— > cut off
 19.5–21.6 **S3** [interruption] But even the workaround last time failed— > takes floor from S1
 21.7–26.8 **S3** [turn] We ended up with patients getting moved three times in one week. That just isn't sustainable. > yields to S4
 24.7–25.4 **S2** [backchannel] Right, right.
 27.0–33.2 **S4** [turn] Not to mention, reception is drowning in calls about it. Yesterday we had ninety-seven voicemails. > yields to S1
 33.1–35.3 **S1** [turn] That's a lot, but if we open up Saturdays— > cut off
 33.8–34.2 **S3** [backchannel] Mm.
 34.5–38.6 **S2** [interruption] We can't open Saturdays! Dr. Patel's contract doesn't allow for it. > takes floor from S1
 38.7–43.1 **S2** [turn] Not unless we pay overtime, which isn't in the budget, and HR will push back. > yields to S4
 43.1–47.1 **S4** [turn] We could ask about comp days, but nobody liked that last quarter. > yields to S1
 47.1–52.8 **S1** [turn] Okay, so, what about telehealth slots for the routine check-ins? That could clear up space for in-person visits.
 50.9–51.4 **S3** [failed claim] Wait—
 52.8–56.7 **S1** [turn] —If we switch even a third of those, it opens at least a dozen slots a week. > yields to S3
 56.6–61.7 **S3** [turn] But patients over sixty hate telehealth. We tried that in January, and more than half refused. > yields to S4
 59.3–59.9 **S2** [backchannel] Correct.
 61.9–67.7 **S4** [turn] We could, I don't know, require telehealth for follow-ups and just make case-by-case exceptions? > yields to S2
 67.7–74.8 **S2** [turn] No, because that's not compliant—we can't force patients off the modality they're comfortable with. It has to be their choice. > yields to S1
 75.0–80.4 **S1** [turn] Alright, so what if we do a callback system? Like, patients on the waitlist get called the moment there's a cancellation. > yields to S3
 80.3–83.6 **S3** [turn] That already happens, but the problem is half the time nobody answers. > yields to S1
 81.4–81.9 **S4** [backchannel] Yeah.
 82.0–82.6 **S2** [backchannel] Right.
 83.5–88.3 **S1** [turn] Sure, but what if we use texting for those notifications instead? With a link to claim the slot?
 86.3–87.2 **S2** [backchannel] That might work.
 86.4–88.4 **S3** [failed claim] Only if the number on file is current—
 88.4–91.5 **S1** [turn] —We can ask at check-in to confirm that, it's not a huge lift. > yields to S4
 92.0–98.1 **S4** [turn] If it's that simple, it's worth a shot. Better than answering the same angry message over and over. > yields to S2

94.2–95.1 **S3** [backchannel] Hm, maybe.
98.7–102.5 **S2** [turn] Let's draft a plan and present it to Dr. Patel by Friday, then. ▷ yields to S1
102.4–104.3 **S1** [turn] Sounds good to me. Thanks, everyone. ▷ yields, end



Transcript of dev_00041

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–5.1 **S1** [turn] Okay, so, I just think standardized tests aren't a fair measure of intelligence at all, like, they don't even capture how creative someone is or how hard they work. ▷ cut off

4.6–9.8 **S4** [interruption] But grades are subjective, standardized tests bring at least some consistency, don't you think? ▷ takes floor from S1 ▷ holds, pauses

10.9–14.6 **S4** [turn] Like, otherwise it's all about who your teacher is and how much they like you.

13.5–14.2 **S3** [backchannel] right

13.8–18.0 **S2** [failed claim] Except, honestly, test anxiety is real and it skews results—

14.7–18.2 **S4** [turn] —and don't even get me started on grade inflation, that's a whole thing. ▷ cut off

17.6–18.6 **S1** [interruption] Well no but— ▷ takes floor from S4

18.7–23.4 **S1** [turn] I mean, it's not like grades are automatically unfair, teachers use rubrics, and— ▷ holds, pauses

20.5–22.1 **S3** [failed claim] That's idealized though.

24.1–26.5 **S1** [turn] —they have professional standards, there's oversight— ▷ yields to S3

26.4–33.0 **S3** [turn] Not always successfully. Plus, if you look at the data, standardized tests correlate with socioeconomic status way more than actual ability, so, like, it's not meritocratic either. ▷ cut off

32.1–35.6 **S4** [interruption] So what's your alternative, just interviews? That's even more biased. ▷ takes floor from S3

35.7–39.6 **S4** [turn] Like every single system has flaws. You can't just throw out testing. ▷ yields to S1

35.7–36.2 **S2** [backchannel] yeah

39.8–44.4 **S1** [turn] Why not have portfolios? Or, I dunno, group projects that show actual skills? ▷ yields to S2

44.3–48.6 **S2** [turn] But then you get freeloaders and it's totally unscalable for big schools. ▷ yields to S3

49.0–54.2 **S3** [turn] Maybe some sort of mixed model? But then it still privileges people with resources. ▷ yields to S4

50.7–51.2 **S1** [backchannel] mm-hm

54.3–60.3 **S4** [turn] I mean, nothing is perfect, but at least exams have a level playing field. That's why they're still used. ▷ yields to S2

60.3–64.1 **S2** [turn] Except for everyone who can't afford prep classes, but, sure. ▷ yields to S3

63.1–63.9 **S1** [backchannel] exactly

64.2–68.8 **S3** [turn] Yeah, that's the core problem. The supposed objectivity is just a veneer. ▷ yields to S1

67.1–67.5 **S4** [laughter] haha

69.4–73.6 **S1** [turn] So, like, does anyone really think SATs test actual intelligence? 'Cause I don't. ▷ yields to S3

70.6–71.1 **S2** [backchannel] nope

73.6–78.2 **S3** [turn] It measures test-taking skills, mostly. And maybe your breakfast that day. ▷ yields to S4

78.4–84.7 **S4** [turn] Alright, but, if you get rid of them, how do you scale admissions at all? Like, it's not even logistically possible. ▷ yields to S2

84.6–87.6 **S2** [turn] Maybe AI can help in the future, honestly. ▷ yields to S3

85.3–85.8 **S1** [backchannel] yeah

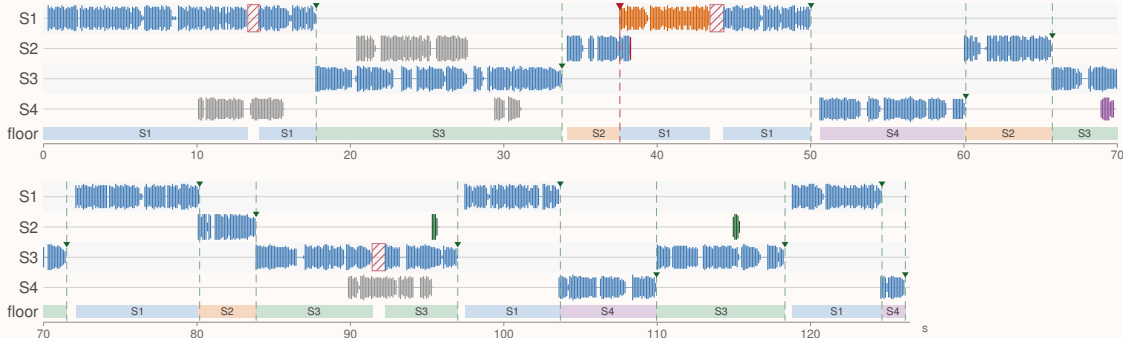
87.6–91.1 **S3** [turn] But then who's programming the AI? That's just new bias. ▷ yields to S4

91.1–96.6 **S4** [turn] True, but status quo's still less risky than, like, total chaos. We need *some* metric. ▷ yields to S1

96.7–99.1 **S1** [turn] We need better metrics, that's all I'm saying. ▷ yields to S3

99.1–101.7 **S3** [turn] Agreed. But the system drags its feet. ▷ yields, end

test split · tone: thoughtful and contemplative · noise: dorm room, 34.7 dB SNR; keyboard clatter · voices: S1 F, S2 M, S3 F, S4 F
EOT 14 yields/3 holds · INT 1 interruptions/3 negatives (2 backchannels, 1 laughter) · 4 failed claims (excluded)



Transcript of test_00003

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–13.3 **S1** [turn] Okay, before we go in circles again, let's just lay out the options for how we could schedule the Wednesday appointments so everyone gets a fair shot. I really think rotating slots is the only way it can work for all the physicians. ▷ holds, pauses

10.1–15.7 **S4** [failed claim] But, like, didn't we already try a rotation last quarter? People just swapped at the last minute anyway—

14.0–17.8 **S1** [turn] —yes, but that was without ground rules. If we had a sign-up cutoff— ▷ yields to S3

17.8–26.2 **S3** [turn] I'd argue enforcing a cutoff is pointless unless you actually penalize late swaps. Otherwise, it just creates drama with no teeth.

20.4–27.7 **S2** [failed claim] Can we pause for a second? The main issue is making sure coverage isn't disrupted, right? I don't think penalties are necessary if we—

26.3–33.8 **S3** [turn] —but that's what keeps happening! Last month, half the patients were rescheduled because two doctors swapped the same day with zero notice. ▷ yields to S2

29.4–31.1 **S4** [failed claim] Oh totally, that was a nightmare.

34.1–38.3 **S2** [turn] Right, I remember that. Still, if we all agree to give at least 48 hours' notice— ▷ cut off

37.6–43.4 **S1** [interruption] But what if someone's kid gets sick? Life happens, and sometimes last-minute swaps are the only option. ▷ takes floor from S2 ▷ holds, pauses

44.3–50.0 **S1** [turn] We can't be so rigid that people feel trapped, but at the same time, patients deserve consistency. ▷ yields to S4

50.6–60.1 **S4** [turn] Honestly, it's all moot if the calendar is a mess. Half the time, there are double-bookings because someone edits the spreadsheet and doesn't tell anyone. I'm just saying! ▷ yields to S2

60.0–65.8 **S2** [turn] That's a fair point. Maybe a shared calendar app with auto-notifications would help instead of the spreadsheet? ▷ yields to S3

65.7–71.5 **S3** [turn] Sure, but tech only goes so far. If people ignore the notifications, we're back at square one. ▷ yields to S1

68.9–69.9 **S4** [laughter] haha exactly

72.1–80.2 **S1** [turn] We could set up an automatic reminder that pings everyone if there's a change. Maybe paired with a quick group confirmation, so no one misses it? ▷ yields to S2

80.1–83.9 **S2** [turn] I... I like that, as long as it's not, you know, overwhelming with messages. ▷ yields to S3

83.9–91.4 **S3** [turn] I still think we're dodging the accountability piece. We need some consequence for last-minute swaps unless it's a true emergency. ▷ holds, pauses

89.9–95.4 **S4** [failed claim] Real talk, who's actually tracking emergencies versus just 'oops, I forgot'? Because that's—

92.3–97.0 **S3** [turn] —that's my point. Without tracking, it's all on the honor system, and people game it. ▷ yields to S1

95.3–95.8 **S2** [backchannel] mm-hm

97.5–103.7 **S1** [turn] Maybe a quarterly review? We could just look at the swap log and see if anyone's abusing it—gently, of course. ▷ yields to S4

103.6–110.0 **S4** [turn] Honestly, I'm fine with that as long as it's not, you know, like a public shaming thing. I think just talking it out is enough. ▷ yields to S3

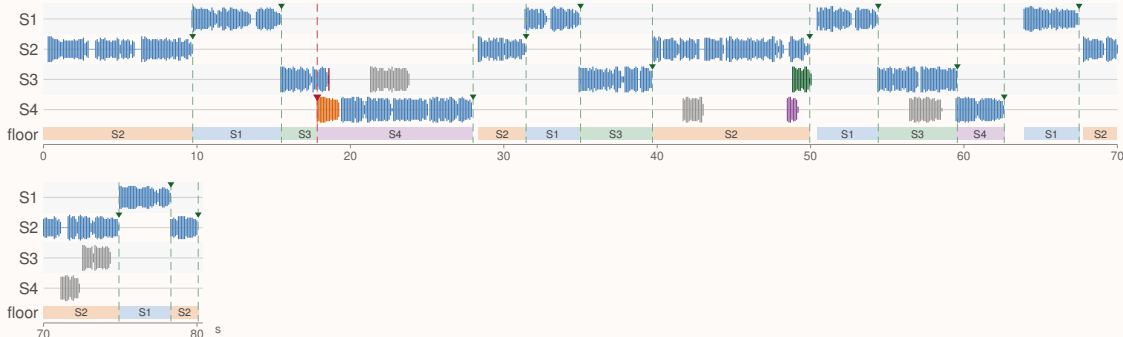
110.0–118.3 **S3** [turn] Alright, so, calendar with notifications, cutoff for swaps except emergencies, and then a review system? I can live with that. ▷ yields to S1

115.0–115.4 **S2** [backchannel] yeah

118.8–124.7 **S1** [turn] That sounds closer to fair, at least. If everyone agrees, maybe we can finally stop arguing about this. ▷ yields to S4

124.6–126.2 **S4** [turn] Big if! But yeah, let's try it. ▷ yields, end

test split · tone: casual and friendly · noise: studio room tone, 32.3 dB SNR; chairs scraping, phone ring · voices: S1 F, S2 M, S3 F, S4 M
EOT 14 yields/0 holds · INT 1 interruptions/2 negatives (1 backchannels, 1 laughter) · 5 failed claims (excluded)



Transcript of test_00051

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.7 **S2** [turn] Alright, so we really need to just pick a place. We've been circling this for like twenty minutes. Can everyone just say what they want so we can vote or something? ▷ yields to S1

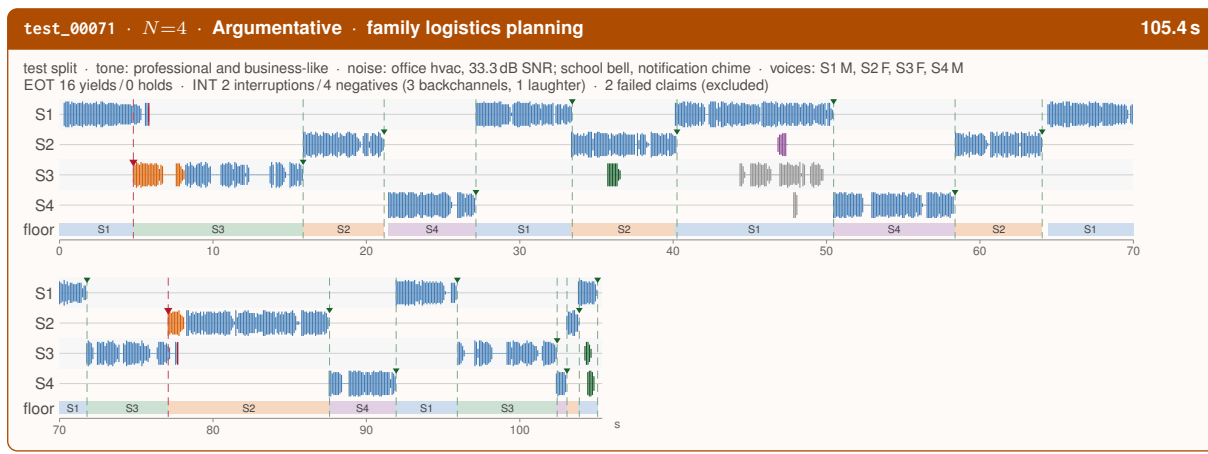
9.7–15.5 **S1** [turn] Thai food makes the most sense, it's close and everyone can find something. I don't get what's so hard about this. ▷ yields to S3

15.5–18.6 **S3** [turn] But, um, didn't we have Thai last week? I just feel like— ▷ cut off

17.8–19.3 **S4** [interruption] Oh come on, last week was sushi— ▷ takes floor from S3

19.4–25.1 **S4** [turn] We had sushi last Friday. Thai was like, what, a month ago? I could totally go for pad see ew right now.

21.3–23.9 **S3** [failed claim] No, I swear it was Thai when Lauren visited—
 25.2–28.0 **S4** [turn] —which was like, way before finals. I'm right on this one. $\triangleright$ yields to S2
 28.3–31.5 **S2** [turn] Guys, can we not get bogged down in calendar trivia? $\triangleright$ yields to S1
 31.4–35.0 **S1** [turn] Honestly, it's just food. We all like Thai, it's not a big deal. $\triangleright$ yields to S3
 34.9–39.7 **S3** [turn] Um, I mean, except I literally can't eat peanuts. Their menu's like, covered in it. $\triangleright$ yields to S2
 39.7–44.4 **S2** [turn] Right, right, peanut allergy, thank you for reminding us, that's actually important.
 41.7–43.1 **S4** [failed claim] Okay, but you've eaten there before—
 44.5–49.9 **S2** [turn] —she always orders the one salad with no nuts and is stuck with, like, lettuce and sadness. $\triangleright$ yields to S1
 48.5–49.2 **S4** [laughter] Haha, true.
 48.8–50.1 **S3** [backchannel] Yeah, not ideal.
 50.4–54.4 **S1** [turn] Okay, so what's your actual suggestion then? Because we're running out of time. $\triangleright$ yields to S3
 54.4–57.1 **S3** [turn] I mean, Italian? Way more options for everyone.
 56.5–58.6 **S4** [failed claim] Ugh, pasta again? I had that for lunch, though—
 57.2–59.6 **S3** [turn] —there's like, risotto or pizza or— $\triangleright$ yields to S4
 59.5–62.6 **S4** [turn] Yeah, but it's still like, heavy? We'll just end up in a carb coma. $\triangleright$ yields to S1
 63.9–67.5 **S1** [turn] Oh my god, just pick something, I can't with this endless debate. $\triangleright$ yields to S2
 67.8–74.9 **S2** [turn] Okay, here's an idea, what about the Mediterranean place? It's new, lots of options, no allergens to worry about. $\triangleright$ yields to S1
 71.1–72.4 **S4** [failed claim] Yeah, I'm down for falafel.
 72.5–74.4 **S3** [failed claim] I'd be good with that, honestly.
 75.0–78.3 **S1** [turn] Fine. If nobody argues for two straight minutes, we order there. $\triangleright$ yields to S2
 78.3–80.1 **S2** [turn] Deal. I'm starting the timer now. $\triangleright$ yields, end

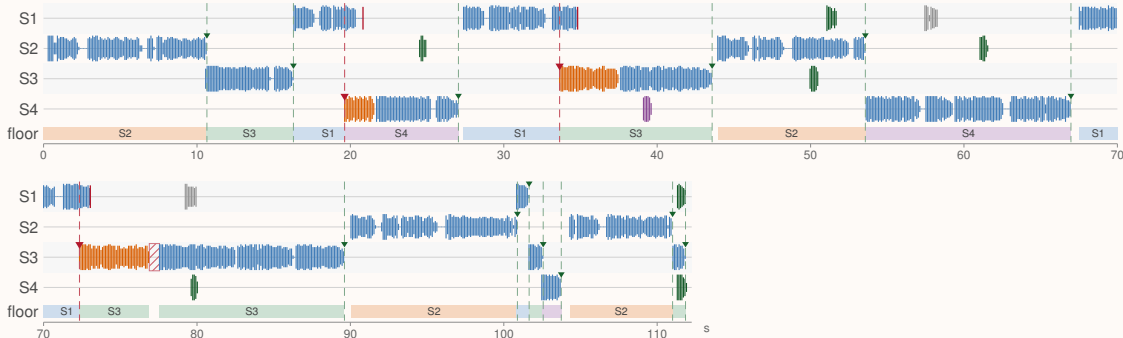


Transcript of test_00071

start—end (s) speaker [ground-truth label] text $\triangleright$ outcome

0.3–5.8 **S1** [turn] Alright, so I think we can just do pizza after soccer, everyone likes pizza, and it's easy. I don't see why we need to overcomplicate this. $\triangleright$ cut off
 4.8–8.1 **S3** [interruption] Except Dad can't eat cheese. Again. $\triangleright$ takes floor from S1
 8.2–15.9 **S3** [turn] And you also seem to forget there's the choir thing at six. That totally overlaps. $\triangleright$ yields to S2
 15.9–21.2 **S2** [turn] Right—and actually, pizza always winds up meaning salad for Mom. She's not thrilled. Ben. $\triangleright$ yields to S4
 21.5–27.1 **S4** [turn] Uh, I could just make something quick at home, honestly, if that's easier for everyone. I don't need pizza either. $\triangleright$ yields to S1
 27.2–33.4 **S1** [turn] I mean, look, it's not just about what you eat, it's about everyone actually getting fed before 8 p.m. for once. $\triangleright$ yields to S2
 33.4–40.3 **S2** [turn] But if you want people to show up for the family meal, they need to be able to actually eat the food. Otherwise what is the point? $\triangleright$ yields to S1
 35.7–36.6 **S3** [backchannel] mmmm
 40.1–45.7 **S1** [turn] But if we start cooking, we're stuck here. We'll be late for choir and soccer runs long every single week—
 44.4–49.8 **S3** [failed claim] No, it only runs late because you chat with the coach for half an hour—
 45.8–50.5 **S1** [turn] —no, it runs late because kids don't get off the field on time, which is not my fault. $\triangleright$ yields to S4
 46.8–47.5 **S2** [laughter] haha
 47.9–48.1 **S4** [failed claim] I—
 50.5–58.4 **S4** [turn] I think we're wandering from solutions a little. Why not just do two quick meals, like wraps for the choir people, something else after soccer for the rest? $\triangleright$ yields to S2
 58.4–64.1 **S2** [turn] Okay, but who's making two meals? Nobody has time, that's why we end up arguing every Sunday. $\triangleright$ yields to S1
 64.4–71.8 **S1** [turn] Exactly. If we just pick up pizza and, like, a couple salads, everyone could at least eat something. It's not perfect, but at least it works. $\triangleright$ yields to S3
 71.8–77.7 **S3** [turn] Sure, let's just initiate the annual pizza versus logistics meltdown. That's a family tradition at this point. $\triangleright$ cut off
 77.1–78.1 **S2** [interruption] Oh, come on— $\triangleright$ takes floor from S3
 78.3–87.6 **S2** [turn] Look, we just need to agree nobody will be totally happy, but people will get actual food and make it to their commitments on time. Can we all live with that for one week? $\triangleright$ yields to S4
 87.6–91.9 **S4** [turn] That's reasonable. I can grab a sandwich before choir and not make it a big deal. $\triangleright$ yields to S1
 91.9–95.9 **S1** [turn] Fine, as long as I don't have to cook after running around all day. Deal? $\triangleright$ yields to S3
 96.0–102.4 **S3** [turn] Deal. I'll order the pizza, because otherwise we'll discuss it until Tuesday. $\triangleright$ yields to S4
 102.4–103.1 **S4** [turn] Thank you. $\triangleright$ yields to S2
 103.1–103.9 **S2** [turn] Perfect. $\triangleright$ yields to S1
 103.8–105.1 **S1** [turn] Alright, meeting adjourned then? $\triangleright$ yields, end
 104.2–104.7 **S3** [backchannel] mm-hm
 104.4–104.9 **S4** [backchannel] Yeah.

test split · tone: thoughtful and contemplative · noise: cafe, 33.2 dB SNR; notification chime, school bell · voices: S1 F, S2 M, S3 M, S4 M
EOT 13 yields / 1 holds · INT 3 interruptions / 8 negatives (7 backchannels, 1 laughter) · 2 failed claims (excluded)



Transcript of test_00119

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–10.7 **S2** [turn] Alright, thanks everyone for joining. We need to lock down the appointment scheduling process for the clinic, because last week's double-booking fiasco can't happen again. ▷ yields to S3

10.6–16.3 **S3** [turn] Totally, but honestly, people just need to check the calendar before picking a slot. Not rocket science. ▷ yields to S1

16.3–20.8 **S1** [turn] That's too simplistic. People forget, or systems don't sync. We need some kind of approval step in there— ▷ cut off

19.6–21.6 **S4** [interruption] But an approval step slows things down— ▷ takes floor from S1

21.7–27.0 **S4** [turn] If someone has to wait for an email every time they need a check-up slot, we'll just make bottlenecks. ▷ yields to S1

24.5–25.0 **S2** [backchannel] right

27.4–34.8 **S1** [turn] It doesn't have to be email—it could be an automated approval that flags just the conflicts. We don't need to approve every single thing by hand. ▷ cut off

33.6–37.5 **S3** [interruption] But then you still get a ton of notifications, and who's actually reading them? ▷ takes floor from S1

37.6–43.6 **S3** [turn] Like, last time IT set up those alerts, my inbox got spammed so bad I just missed all the actual issues. ▷ yields to S2

39.1–39.7 **S4** [laughter] haha

43.9–53.6 **S2** [turn] But is volume really the problem, or are we talking about clarity? Because if notifications only go out when there's a real conflict, wouldn't that work? ▷ yields to S4

50.0–50.5 **S3** [backchannel] mm, maybe

51.1–51.8 **S1** [backchannel] Exactly.

53.6–59.3 **S4** [turn] But that still depends on the system catching every possible conflict, and we've already had sync issues—

57.4–58.3 **S1** [failed claim] Right, except—

59.4–67.0 **S4** [turn] —and we can't trust that just updating the app will make that go away. So unless IT guarantees it's ironclad, human review seems safer. ▷ yields to S1

61.0–61.6 **S2** [backchannel] uh-huh

67.5–73.0 **S1** [turn] But human review is precisely what leads to our delays. If we automate the double-booking detection, then only edge cases get flagged. Most cases would be smooth. ▷ cut off

72.3–76.9 **S3** [interruption] But what about last Tuesday? We had software in place, and it still missed that morning slot. ▷ takes floor from S1 ▷ holds, pauses

77.6–82.5 **S3** [turn] Like, it's not just about delays, it's about actually not screwing up someone's test schedule, you know?

79.2–80.0 **S1** [failed claim] Okay, but—

79.6–80.1 **S4** [backchannel] yeah

82.7–89.6 **S3** [turn] —if one patient gets bumped they can't always reschedule for weeks. We're not just talking admin headaches, it's patient outcomes. ▷ yields to S2

90.0–100.9 **S2** [turn] So what if we run a hybrid—software first, but with a human fallback for exceptions? Would both sides be satisfied if there's a daily manual review of conflicts? ▷ yields to S1

100.8–101.7 **S1** [turn] That's workable. ▷ yields to S3

101.6–102.6 **S3** [turn] Yeah, that could fly. ▷ yields to S4

102.5–103.7 **S4** [turn] I could get behind that. ▷ yields to S2

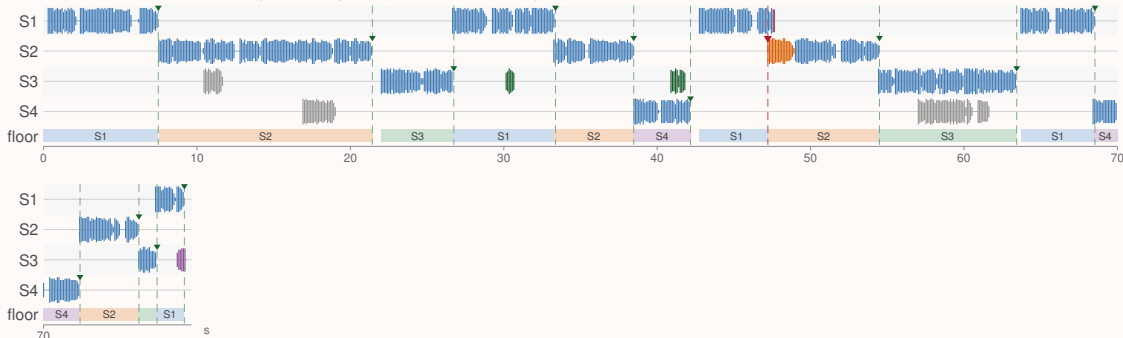
104.3–111.0 **S2** [turn] Great, I'll draft a new workflow—can someone from IT check the sync issue before we roll anything out, though? ▷ yields to S3

111.0–111.8 **S3** [turn] Yeah, I'll ping them. ▷ yields, end

111.3–112.0 **S4** [backchannel] Perfect.

111.3–111.9 **S1** [backchannel] Thank you.

test split · tone: professional and business-like · noise: office hum, 30.9 dB SNR; alarm beep · voices: S1 M, S2 F, S3 F, S4 M
EOT 13 yields / 0 holds · INT 1 interruptions / 3 negatives (2 backchannels, 1 laughter) · 3 failed claims (excluded)



Transcript of test_00125

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–7.5 **S1** [turn] All right, before we get too sidetracked, can we just zero in on where everyone actually wants to eat tonight? We are burning time here. ▷ yields to S2

7.5–14.2 **S2** [turn] I'm just going to say it: if we go to Giuseppe's again, I'm out. It's overpriced and the service is slow. Let's pick somewhere efficient.

10.5–11.7 **S3** [failed claim] But some people—

14.2–18.9 **S2** [turn] —and honestly, their vegetarian options are basically just salad with extra bread.

16.9–19.1 **S4** [failed claim] Ah, but uh, the Italian place does have—

19.0–21.4 **S2** [turn] —and the wait times are never under an hour. $\triangleright$ yields to S3

22.0–26.7 **S3** [turn] Okay, but like, some of us aren't just looking for speed. The atmosphere matters too, you know? $\triangleright$ yields to S1

26.7–33.4 **S1** [turn] That's fair, but there has to be a middle ground. What about the Korean place? It's close, fast, everyone finds something there. $\triangleright$ yields to S2

30.2–30.7 **S3** [backchannel] mm-hm

33.3–38.5 **S2** [turn] Not happening. The last time we went there, half of us left still hungry because the portions are tiny. $\triangleright$ yields to S4

38.5–42.2 **S4** [turn] But, uh, their kimchi pancakes—those were good, and it's quiet enough to talk. $\triangleright$ yields to S1

40.9–41.9 **S3** [backchannel] Exactly.

42.8–47.6 **S1** [turn] Nobody's forcing anyone to go where they don't want, but we need a list. What about the tapas place on 5th? $\triangleright$ cut off

47.2–48.9 **S2** [interruption] No way, the bill last time— $\triangleright$ takes floor from S1

49.0–54.5 **S2** [turn] It just spirals out of control with small plates. I don't want another argument about who ordered what. $\triangleright$ yields to S3

54.4–60.0 **S3** [turn] But it's so good. And we could just agree to split the check evenly, right? That worked at the sushi place.

57.0–61.7 **S4** [failed claim] I mean... I'd be fine with that, as long as we set the expectation up front. Maybe we just—

60.0–63.4 **S3** [turn] —as long as we don't order a million things nobody touches. $\triangleright$ yields to S1

63.7–68.5 **S1** [turn] So it's down to tapas or Korean? Can we just vote or do we need to rehash this again? $\triangleright$ yields to S4

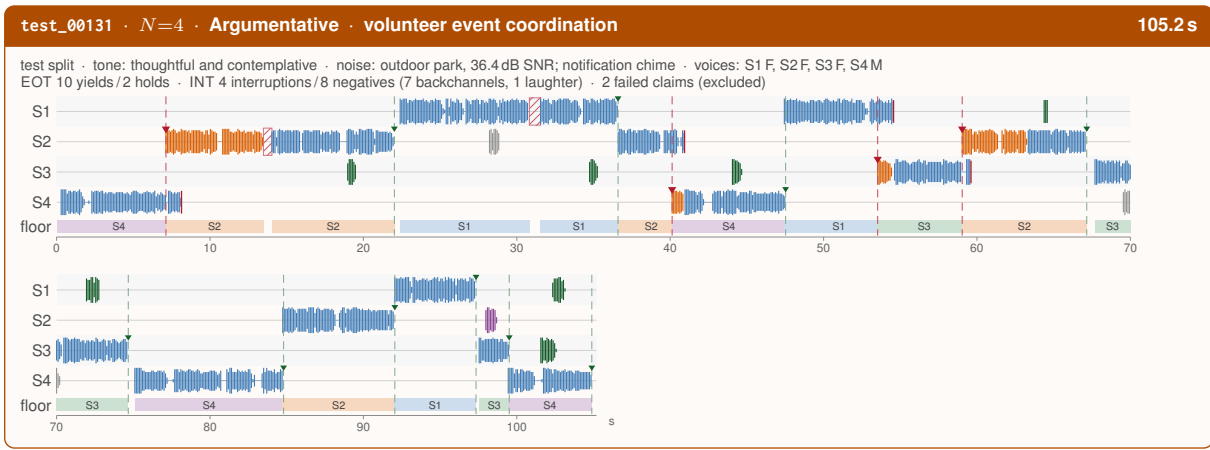
68.4–72.4 **S4** [turn] Uh, I think... yeah, let's just vote. I'm leaning tapas if we agree to the split. $\triangleright$ yields to S2

72.3–76.2 **S2** [turn] Fine, but next time someone else gets to pick first. Tapas then. $\triangleright$ yields to S3

76.2–77.4 **S3** [turn] Yeah, that works. $\triangleright$ yields to S1

77.3–79.2 **S1** [turn] All right, decision made. Finally. $\triangleright$ yields, end

78.7–79.3 **S3** [laughter] haha



Transcript of test_00131

start–end (s) speaker [ground-truth label] text $\triangleright$ outcome

0.3–8.1 **S4** [turn] Alright, let's get started. The main issue on the table is whether we move the volunteer event to next weekend or keep it as scheduled. Who wants to— $\triangleright$ cut off

7.1–13.5 **S2** [interruption] Me! Honestly, I really think moving it is a mistake, I mean, all our outreach is already built around this Saturday. $\triangleright$ takes floor from S4 $\triangleright$ holds, pauses

14.0–22.0 **S2** [turn] We have flyers, emails sent out—changing it now would just confuse everybody. Besides, we finally got all the food vendors lined up. $\triangleright$ yields to S1

19.0–19.6 **S3** [backchannel] mm-hm

22.4–30.8 **S1** [turn] But, um, isn't the weather supposed to be really bad? Like, thunderstorms. I just worry about putting people at risk, especially the older volunteers— $\triangleright$ holds, pauses

28.2–28.9 **S2** [failed claim] That's not—

31.6–36.6 **S1** [turn] —and some of them can't drive in that kind of storm. If we postpone, it's safer for everyone. $\triangleright$ yields to S2

34.7–35.3 **S3** [backchannel] right

36.6–40.9 **S2** [turn] You guys, the forecast changes all the time, though. What if it clears up? We'll lose all the momentum if we change it for no reason— $\triangleright$ cut off

40.1–40.9 **S4** [interruption] Can I just— $\triangleright$ takes floor from S2

41.0–47.5 **S4** [turn] What about a backup plan? Could we have an indoor option or a rain date that we communicate clearly to everyone? $\triangleright$ yields to S1

44.1–44.7 **S3** [backchannel] um-hm

47.4–54.6 **S1** [turn] That's a good idea, but would we even have enough room inside the community center for all the activities? Last year it was super cramped— $\triangleright$ cut off

53.5–54.5 **S3** [interruption] I just think— $\triangleright$ takes floor from S1

54.6–59.6 **S3** [turn] Sorry, but if it keeps people from getting soaked or slipping in the mud, I think we should consider it. It's not perfect but it's better than nothing. $\triangleright$ cut off

59.0–63.3 **S2** [interruption] But nobody wants to be stuck inside all day, it's not the same energy at all! $\triangleright$ takes floor from S3

63.3–67.2 **S2** [turn] Half the reason people sign up is for the park, the booths, just being outside. $\triangleright$ yields to S3

64.4–64.6 **S1** [backchannel] mm

67.7–70.4 **S3** [turn] I know, but it's also about showing up for each other, right?

69.5–70.2 **S4** [failed claim] Hold on—

70.5–74.7 **S3** [turn] —and making sure everyone goes home safe matters more than perfect photos or the vibe. $\triangleright$ yields to S4

72.0–72.8 **S1** [backchannel] exactly

75.1–84.8 **S4** [turn] Let's try to agree on a contingency: if it's raining Saturday morning, we move it indoors; if not, we go ahead as planned. Can everyone live with that? $\triangleright$ yields to S2

84.7–92.1 **S2** [turn] Fine, I guess, but only if someone else calls the food vendors, because I can't do all the updates again if it changes last-minute. $\triangleright$ yields to S1

92.0–97.3 **S1** [turn] I can help with that. I just really want to make sure no one feels left out or put in a tough spot. $\triangleright$ yields to S3

97.5–99.5 **S3** [turn] Thank you, Clara. That sounds fair to me. $\triangleright$ yields to S4

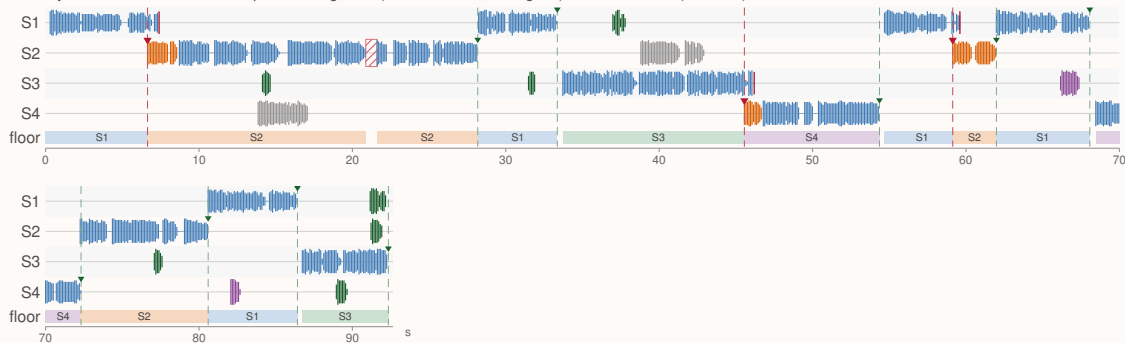
98.0–98.7 **S2** [laughter] haha okay

99.4–104.9 **S4** [turn] Alright, let's call it a plan then. I'll draft a backup announcement tonight and send it for review. $\triangleright$ yields, end

101.6–102.6 **S3** [backchannel] thanks everyone

102.3–103.2 **S1** [backchannel] Perfect.

test split · tone: casual and friendly · noise: office hum, 22.9 dB SNR · voices: S1 M, S2 F, S3 F, S4 M
EOT 9 yields / 1 holds · INT 3 interruptions / 9 negatives (7 backchannels, 2 laughter) · 2 failed claims (excluded)



Transcript of test_00133

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–7.4 **S1** [turn] Alright, so—look, I just want to say we cannot keep kicking the notification redesign down the road. Users are still saying it's, like, impossible to manage their settings, and it's getting embarrassing. ▷ cut off

6.6–8.6 **S2** [interruption] I don't think that's actually true, I mean— ▷ takes floor from S1

8.7–20.9 **S2** [turn] If you look at the last survey data, the complaints are mostly about the language, not the settings screen itself. Isn't it more efficient to focus on clarifying the copy first instead of a whole redesign? ▷ holds, pauses

13.9–17.2 **S4** [failed claim] But, uh, the copy rewrite won't fix the, the—confusing toggles, right?

14.1–14.7 **S3** [backchannel] mm-hm

21.6–28.2 **S2** [turn] —interface. I mean, sure, the toggles could use love, but people are clearly lost on the explanations first. ▷ yields to S1

28.2–33.4 **S1** [turn] No, but wait, Max is totally right, the toggles literally say 'off' but they're green, it makes no sense! ▷ yields to S3

31.5–31.9 **S3** [backchannel] yeah

33.7–41.7 **S3** [turn] I mean, I—I see both sides, but the actual feedback from onboarding is about, like, people just giving up and turning everything off, which is—

37.0–37.8 **S1** [backchannel] Exactly!

38.8–42.9 **S2** [failed claim] Because they don't trust it'll work if they leave some on! That's a copy problem.

41.8–46.2 **S3** [turn] —such a mix of UI and trust, I honestly think we need a prototype, just a quick Figma mock to test both? ▷ cut off

45.6–46.7 **S4** [interruption] Can—can I say— ▷ takes floor from S3

46.8–54.4 **S4** [turn] Um, so, we've already got the old prototype, right? We could just—add new wording to it and see if that helps before, you know, building something completely new. ▷ yields to S1

54.7–59.6 **S1** [turn] But the visual confusion's not gonna change if we just touch up the words, man! Like, the toggle colors are the number one complaint in usertesting. ▷ cut off

59.1–62.0 **S2** [interruption] No, that's anecdotal, show me where it's number one— ▷ takes floor from S1 ▷ yields to S1

62.0–68.1 **S1** [turn] The last round, literally, like, seven out of ten said the toggles were backwards. I can pull the spreadsheet up right now. ▷ yields to S4

66.2–67.4 **S3** [laughter] haha, oh my god

68.5–72.3 **S4** [turn] I, uh, I'd like to actually see that data, if you don't mind emailing it after? ▷ yields to S2

72.3–80.6 **S2** [turn] Okay, but until we see that, can we just agree to test wording and layout at the same time, not sequentially? Otherwise we'll waste weeks. ▷ yields to S1

77.1–77.6 **S3** [backchannel] right

80.6–86.4 **S1** [turn] Fine, but I swear, if people still hate the toggles after the copy fix, we're not arguing about it next sprint. ▷ yields to S3

82.1–82.7 **S4** [laughter] haha, deal

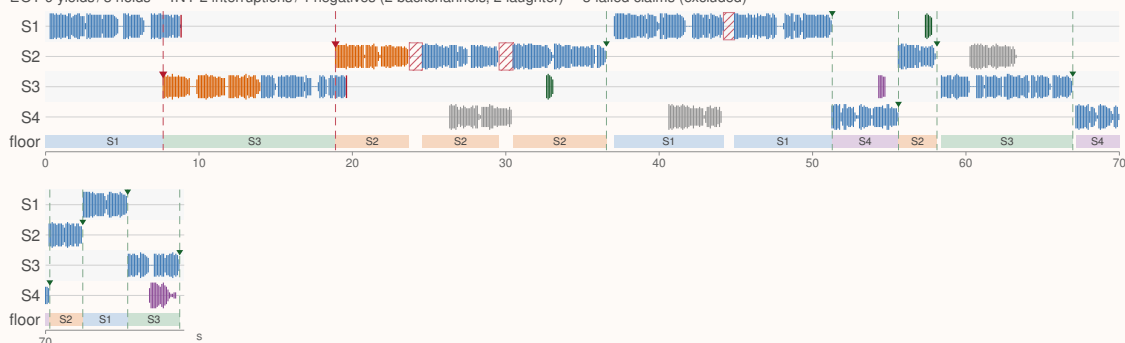
86.7–92.3 **S3** [turn] Alright, so, dual track: copy and layout. I'll spin up a prototype tonight if that works for everyone? ▷ yields, end

89.0–89.7 **S4** [backchannel] Perfect.

91.2–92.2 **S1** [backchannel] Yeah, thanks Jess.

91.2–92.0 **S2** [backchannel] Sounds good.

test split · tone: casual and friendly · noise: living room tv distant, 30.8 dB SNR; keyboard clatter; distant speech · voices: S1 M, S2 M, S3 F, S4 M
EOT 9 yields / 3 holds · INT 2 interruptions / 4 negatives (2 backchannels, 2 laughter) · 3 failed claims (excluded)



Transcript of test_00172

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.8 **S1** [turn] Alright, so for the bonus question, they're asking, uh, what was the original flavor of Gatorade. I'm pretty sure it's lemon-lime. Like, I read about the University of Florida and— ▷ cut off

7.7–14.0 **S3** [interruption] Wait, no, I swear it was orange! My uncle always says it was orange, he was obsessed with that trivia fact. ▷ takes floor from S1

14.1–19.6 **S3** [turn] It makes sense, like, sports drinks were all orange at first. I mean, when you think of Gatorade, the orange one was, like, everywhere. ▷ cut off

18.9–23.7 **S2** [interruption] Uh, nah, it had to be lemon-lime, right? That's the classic. They even called it the green one, man. ▷ takes floor from S3 ▷ holds, pauses

24.6–29.6 **S2** [turn] Like, that's the one you get in those old NFL photos. Orange came later, I'm pretty sure. ▷ holds, pauses

26.3–30.4 **S4** [failed claim] No, no, I remember the commercials, they were always dumping orange on the coach.

30.5–36.6 **S2** [turn] —on the coach, yeah, but that doesn't mean it was first! Advertisers just liked orange 'cause it stands out. Trust me. ▷ yields to S1

32.7–33.1 **S3** [backchannel] right

37.1–44.2 **S1** [turn] So, uh, technically I think both existed pretty early, but the first one made was lemon-lime. They invented it for the football team and— ▷ holds, pauses

40.6–44.1 **S4** [failed claim] Hold on, I still say orange is the OG, I'm not convinced.

44.9–51.3 **S1** [turn] —and that was the only flavor for a couple of years, if I recall. Orange got, um, added later on, maybe '69? ▷ yields to S4

51.3–55.6 **S4** [turn] No way, that feels too late. I was a kid then, and I remember both at the same time. ▷ yields to S2

54.3–54.8 **S3** [laughter] haha

55.6–58.1 **S2** [turn] Dude, that's the Mandela Effect, you're just mixing it up. ▷ yields to S3

57.4–57.8 **S1** [backchannel] mm-hm

58.4–64.1 **S3** [turn] Okay but like, does it actually matter? I mean, this is just a bonus, we're fighting like it's life or death.

60.3–63.3 **S2** [failed claim] You just want to put down orange 'cause it matches your shirt, don't you.

64.3–67.0 **S3** [turn] —hey! That's not true, but now I kinda want some. ▷ yields to S4

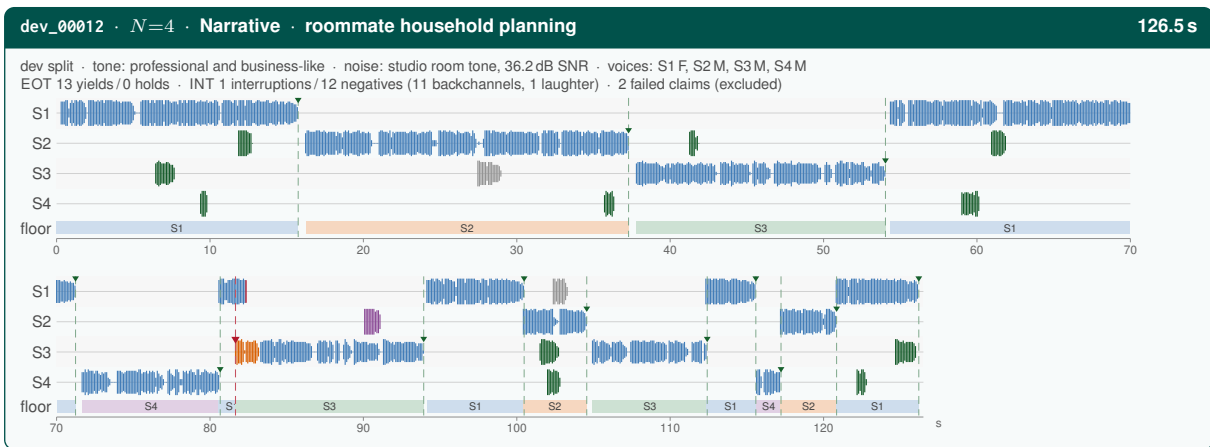
67.2–70.3 **S4** [turn] Alright, what if we just vote? I say orange. Who's with me? ▷ yields to S2

70.2–72.4 **S2** [turn] Lemon-lime, all the way. Gotta stay classic. ▷ yields to S1

72.5–75.4 **S1** [turn] Yeah, lemon-lime for me too. Sorry, orange fans. ▷ yields to S3

75.4–78.7 **S3** [turn] Ohhh, fine! Outvoted. But I'm fact-checking you guys later. ▷ yields, end

76.8–78.5 **S4** [laughter] haha yeah, you'll see!



Transcript of dev_00012

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–15.8 **S1** [turn] All right, let's get started. I'd like to talk about how we're going to handle chores this month. I know last month got a little chaotic, with the trash not going out and dishes left in the sink, so I think it's worth taking a few minutes to come up with a schedule that actually works for everyone. ▷ yields to S2

6.5–7.8 **S3** [backchannel] Yeah exactly.

9.4–9.9 **S4** [backchannel] Sure.

11.9–12.8 **S2** [backchannel] Right, agreed.

16.2–31.5 **S2** [turn] So, I was thinking about when I lived at my old apartment, and we actually had a rotating chart. Everyone just took pictures with their phone on Sunday night and sent them to the group—made it hard to slack off, honestly, since there was proof. Uh, I'm not saying we need to do that, but maybe something a little more systematic would help?

27.5–29.0 **S3** [failed claim] That's not a bad idea—

31.7–37.3 **S2** [turn] —since, like, last week, I think I took out recycling three times in a row and I didn't even realize till Thursday. ▷ yields to S3

35.7–36.4 **S4** [backchannel] Same here.

37.8–54.0 **S3** [turn] I really don't mind doing extra sometimes, but if people forget their week entirely, then it gets old pretty fast. I remember in college, my roommate just... never cleaned the bathroom. We tried leaving notes, texts, even putting her name on the calendar. Nothing. So, yeah, I'd support some kind of tracking, like, even a whiteboard. ▷ yields to S1

41.3–41.9 **S2** [backchannel] Yep.

54.4–71.2 **S1** [turn] That's a good point. I'd be fine with a whiteboard, and I also think if we just set reminders on our phones, that covers most of it. What matters to me is that we've got, like, accountability, but also that if something comes up—say you can't do your turn—that's flagged in advance rather than discovered two days later. ▷ yields to S4

59.0–60.2 **S4** [backchannel] Reminders make sense.

61.0–61.9 **S2** [backchannel] Yeah, agreed.

71.7–80.7 **S4** [turn] I can write up a chart if you want. Or just use the one I did last year at my old place. It's, uh, just a grid with dates and names in sharpie. ▷ yields to S1

80.6–82.3 **S1** [turn] Perfect, if you could bring that for reference— ▷ cut off

81.7–83.2 **S3** [interruption] Wait, wait, can I just say— ▷ takes floor from S1

83.3–93.9 **S3** [turn] If we're doing a chart, can we actually put it somewhere we see it all the time? Because if it's on the fridge—like, trust me, that's when people actually notice. Otherwise I guarantee it gets ignored after a week. ▷ yields to S1

90.1–91.1 **S2** [laughter] Haha, yeah, true.

94.1–100.5 **S1** [turn] That's a good call. The fridge it is. All right, what about bathroom cleaning? That's honestly been the biggest issue lately. ▷ yields to S2

100.4–104.6 **S2** [turn] Uh, I can take the first week, if that's all right. I'll be home most evenings, so it's easy. ▷ yields to S3

101.5–102.8 **S3** [backchannel] Yeah, that's good.

102.0–102.8 **S4** [backchannel] Fine by me.

102.4–103.3 **S1** [failed claim] Works for me too.

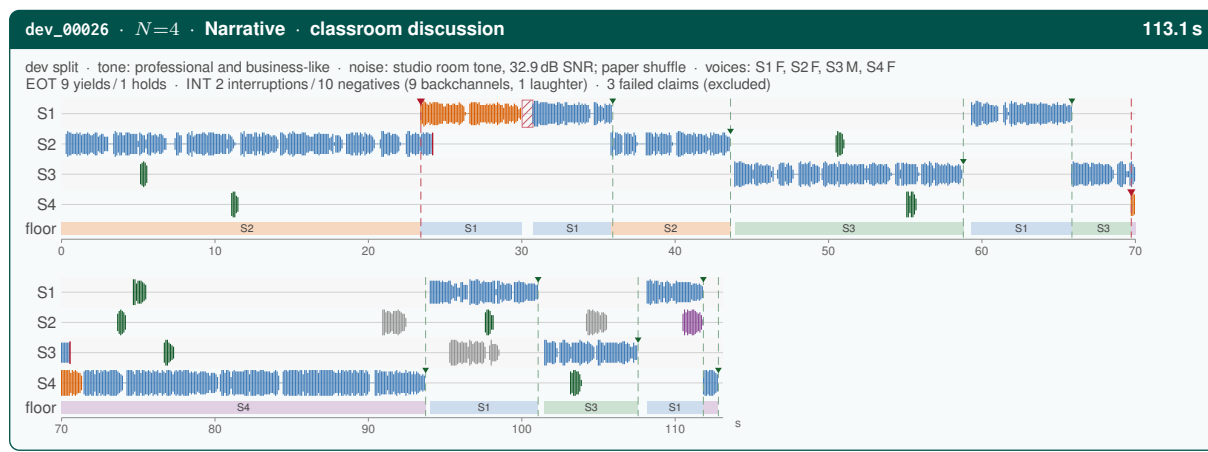
104.9–112.4 **S3** [turn] So I'll take the second week, just for the record. And if someone can cover me the last week of June, I'll be traveling for work. I can swap, but— ▷ yields to S1

112.3–115.6 **S1** [turn] Noted, I can swap in if you remind me a few days before. ▷ yields to S4

115.6–117.2 **S4** [turn] Yeah, I can help if needed. ▷ yields to S2

117.2–120.8 **S2** [turn] Thanks, guys. So, uh, should we just finalize the grid tonight, then? Print it out or— ▷ yields to S1

120.8–126.2 **S1** [turn] I can do that. I'll make it editable and email everyone so if you want to swap a week, it's clear. > yields, end
 122.2–122.8 **S4** [backchannel] Okay.
 124.7–126.0 **S3** [backchannel] Thanks, appreciate it.



Transcript of dev_00026

start–end (s) speaker [ground-truth label] text > outcome

0.3–14.1 **S2** [turn] So I wanted to share something that happened to me last semester, when I was trying to organize a group project in Dr. Chen's class. Um, it was actually really hard to get everyone on the same page, especially because half of us barely knew each other.

5.1–5.6 **S3** [backchannel] Yeah.

11.1–11.6 **S4** [backchannel] Right.

14.3–24.2 **S2** [turn] Um, we ended up meeting at the library, but no one really wanted to take the lead, and we sort of spent the first hour just... talking around the assignment. Eventually I just blurted out, like, "Can someone write down what we're supposed to do?" and that finally got us moving. > cut off

23.4–30.0 **S1** [interruption] Was that because no one wanted to take responsibility, or was it more confusion about the actual project requirements? > takes floor from S2 > holds, pauses

30.7–35.9 **S1** [turn] Because I feel like that happens a lot in classes where the instructions are, honestly, pretty vague. > yields to S2

35.8–43.6 **S2** [turn] Yeah, totally! I think it was both. We didn't know each other, and also, like, the guidelines were just a bullet list online, so everyone was guessing. > yields to S3

43.9–58.8 **S3** [turn] That's why I usually just take charge in those situations. People want direction. Once, in my sophomore year, I actually assigned everyone roles in the first ten minutes, and nobody really objected. It made things run way smoother—even the professor commented on how organized we seemed. > yields to S1

50.5–51.0 **S2** [backchannel] Wow.

55.1–55.8 **S4** [backchannel] Makes sense.

59.3–65.9 **S1** [turn] But do you ever get pushback on that, though? Like, sometimes people resent someone just jumping in and telling everyone what to do. > yields to S3

65.8–70.6 **S3** [turn] Sometimes, but not if you phrase it more like a suggestion. Or ask, 'Hey, can we each take one part?' That way you're still leading, but subtly. > cut off

69.7–71.4 **S4** [interruption] Actually, I want to jump in here— > takes floor from S3

71.5–80.2 **S4** [turn] Because I think there's also a gender element. I've noticed women tend to hesitate more to take charge, and it's not always received the same way if they do.

73.7–74.2 **S2** [backchannel] Yeah.

74.7–75.6 **S1** [backchannel] Absolutely.

76.7–77.3 **S3** [backchannel] Hmm.

80.4–93.7 **S4** [turn] So, my approach is to be very clear and business-like at the start. I just say, 'Okay, here's the agenda, here are the deadlines, does anyone have strong preferences?' That usually diffuses tension and sets expectations. > yields to S1

90.9–92.5 **S2** [failed claim] That sounds so organized.

94.0–96.5 **S1** [turn] I wish I'd done that in my last group, honestly.

95.3–98.5 **S3** [failed claim] You can always change your approach next time. Actually—

96.6–101.1 **S1** [turn] —I think I'm going to try being more upfront next time, instead of waiting for someone else. > yields to S3

97.6–98.2 **S2** [backchannel] Right.

101.5–107.6 **S3** [turn] That's the spirit. And honestly, if you just say what you're thinking, lots of people are relieved someone else said it first. > yields to S1

103.2–103.9 **S4** [backchannel] Exactly.

104.2–105.6 **S2** [failed claim] Yeah, that's so true.

108.2–111.8 **S1** [turn] Alright, so next group project, I guess we'll all be the bossy ones then. > yields to S4

110.5–111.8 **S2** [laughter] Haha, definitely.

111.8–112.8 **S4** [turn] Looking forward to it. > yields, end

test split · tone: professional and business-like · noise: living room tv distant, 24.3dB SNR; dishes clinking · voices: S1 M, S2 M, S3 F, S4 F
EOT 14 yields / 1 holds · INT 1 interruptions / 6 negatives (4 backchannels, 2 laughter) · 3 failed claims (excluded)

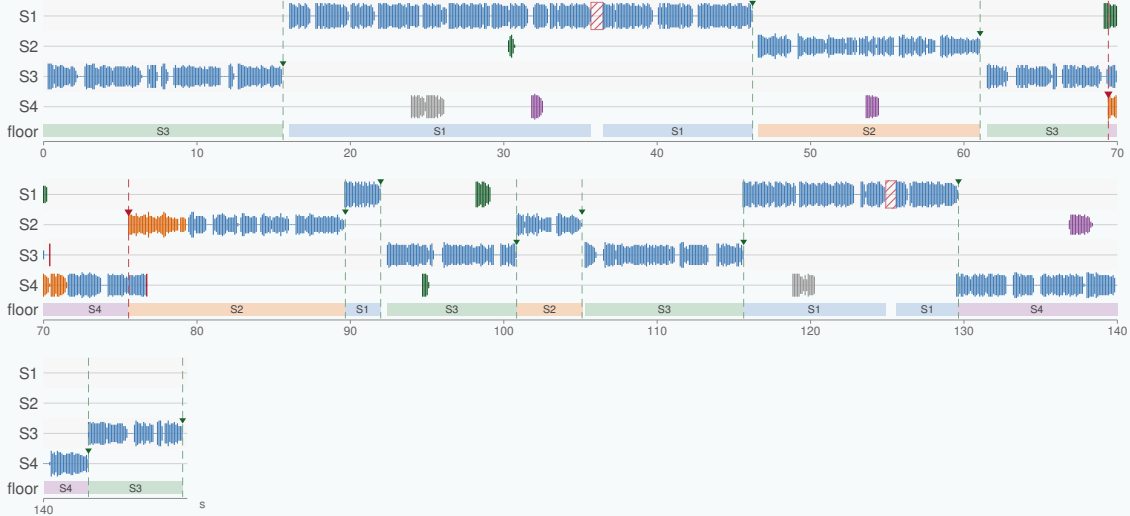


Transcript of test_00025

start—end (s) speaker [ground-truth label] text ▷ outcome

- 0.3–17.5 **S1** [turn] Alright everyone, since we're all here, I thought I'd kick us off with a quick story from last year. You know, that food drive we did at the elementary school? I still remember showing up before dawn—the principal let me in, but none of the tables were even set up, so I just... well, I ended up stacking cans in the dark for half an hour until someone found the light switches. ▷ holds, pauses
- 4.2–5.7 **S2** [backchannel] Yeah, that sounds—wow.
- 8.1–10.8 **S4** [failed claim] Oh no, that sounds like classic volunteer chaos!
- 18.6–29.0 **S1** [turn] Haha yeah, so uh, point being, I'd love if this year we could, you know, actually have a setup team scheduled the night before. Maybe even check if we'll have keys. Just so we're not feeling around in the dark again. ▷ yields to S3
- 29.5–43.9 **S3** [turn] Agreed. I can contact the school and confirm access for Friday evening. If they can't, we'll need a backup. I'd also like to suggest we label donation bins ahead of time—last spring, people kept asking me which box was for pasta and I was trying to run registration. ▷ yields to S4
- 43.8–48.1 **S4** [turn] Oh my gosh, yes! Remember the kid who tried to donate his frog lunchbox?
- 45.9–47.3 **S2** [laughter] Haha, I remember—
- 48.2–53.7 **S4** [turn] He was so proud of it! I had to convince him to keep his lunch. Kids are honestly the best part. ▷ yields to S2
- 54.1–62.2 **S2** [turn] Um, actually, that, that reminds me—last year, uh, one of the parents brought, uh, like five extra volunteers without telling us. And they all just showed up at once and, I, I didn't know who was supposed to do what. So, uh—▷ cut off
- 61.6–66.5 **S3** [interruption] I can set up a sign-in sheet, that way we know exactly who's expected and who's just... there. ▷ takes floor from S2
- 66.6–73.2 **S3** [turn] I'll bring clipboards and have name tags ready, so anyone extra gets added on the spot. No more mystery helpers. ▷ yields to S1
- 73.6–79.1 **S1** [turn] Good call. Last thing we need is, uh, people wandering around without any idea what they're supposed to do. ▷ yields to S3
- 77.0–77.8 **S4** [backchannel] Right, right.
- 77.6–78.5 **S2** [backchannel] Exactly.
- 79.2–88.0 **S3** [turn] I'll also make a tent with colored zones, so we can say, 'If you're in orange, you're food sorting, green you're packing, yellow you run carts to cars.' It'll help with the rush.
- 83.8–85.4 **S4** [failed claim] Wait—oh, sorry, can I jump in?
- 88.2–91.8 **S3** [turn] —and I'll have the zones up by the entrance so no one misses them. ▷ yields to S4
- 91.8–107.5 **S4** [turn] Okay, thanks! So—what if we do a little 'welcome huddle' too, right at the start? Just, like, five minutes. It helps everyone relax, and we can do a fun icebreaker or something. It really changed the mood at the last bake sale; everyone suddenly knew each other's names. ▷ yields to S1
- 107.9–112.7 **S1** [turn] I like that. People remember the orientation better if you can make 'em laugh, even once. ▷ yields to S2
- 109.8–111.3 **S2** [failed claim] That, uh, that makes sense.
- 112.7–120.7 **S2** [turn] Could I also... um, just bring up the lunch situation? Last year, people got sort of lost and, and didn't know when to take breaks. Should we set scheduled shifts? ▷ yields to S1
- 121.3–128.3 **S1** [turn] Yeah, definitely—a schedule for breaks is a must. Otherwise by noon half the crew's disappeared and the rest are hungry and cranky. ▷ yields to S3
- 126.5–127.4 **S4** [laughter] Haha, guilty!
- 128.1–134.2 **S3** [turn] Alright, I'll draft a master plan. And let's circulate it by Wednesday so everyone has time to ask questions? ▷ yields to S1
- 131.4–132.5 **S2** [backchannel] Sounds good.
- 134.5–137.8 **S1** [turn] Thank you, everyone. This is already running smoother than last year. ▷ yields to S4
- 137.7–141.6 **S4** [turn] Here's hoping none of us get stuck in a closet with canned beans again, huh? ▷ yields to S3
- 141.6–142.8 **S3** [turn] Let's make sure of it. ▷ yields, end

test split · tone: curious and inquisitive · noise: office hvac, 36.7 dB SNR; laughter offmic, paper shuffle · voices: S1 M, S2 M, S3 M, S4 F
EOT 11 yields / 2 holds · INT 2 interruptions / 7 negatives (4 backchannels, 3 laughter) · 2 failed claims (excluded)



Transcript of test_00042

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–15.6 **S3** [turn] Alright, thanks everyone for joining. So, I thought we'd kick off planning for the river cleanup next Saturday, but first, um, maybe we could share a story or two from past events? Like, what's something memorable that's happened to you volunteering? ▷ yields to S1

16.0–26.3 **S1** [turn] Uh, I could start, I guess. Um, so the first time I ever did one of these—I think it was like two years ago?—I was so nervous that I'd, uh, mess up somehow, y'know? But then, this lady came up and—

24.0–26.2 **S4** [failed claim] Oh let me guess, she gave you a bag that broke?

26.4–35.7 **S1** [turn] —no, actually! She handed me a pair of those giant tongs, but then immediately, like, tripped right in front of me. I totally panicked, almost dropped the tongs, but she just burst out laughing. ▷ holds, pauses

30.3–30.8 **S2** [backchannel] Mm-hm

31.8–32.6 **S4** [laughter] Haha no way

36.5–46.2 **S1** [turn] Yeah! And she said, 'Don't worry, everyone fumbles on their first time.' It, uh, really made me relax after that. I still think about it whenever I'm, you know, nervous about organizing stuff. ▷ yields to S2

46.6–61.1 **S2** [turn] That's a good one. I remember something sort of similar. Back in, oh, maybe 2015, I was helping with a food drive, not rivers, but—I, uh, mixed up the donation boxes. Ended up putting canned peaches in the pet food bin. Lady running the table didn't let me forget it for weeks. ▷ yields to S3

53.6–54.5 **S4** [laughter] Haha classic

61.5–70.4 **S3** [turn] Oh, I totally get that. My first park cleanup, I managed to, uh, collect almost only bottle caps for like, half an hour. Didn't realize there was a contest going for the weirdest trash find until someone brought over a garden gnome. ▷ cut off

69.1–70.2 **S1** [backchannel] A garden gnome?

69.4–71.5 **S4** [interruption] Wait, seriously, like a full-size gnome? ▷ takes floor from S3

71.6–76.7 **S4** [turn] Because that beats anything I ever found, I just get weird socks or the occasional rubber duck. ▷ cut off

75.5–79.4 **S2** [interruption] Oh, you haven't heard about the time we pulled a bicycle out of the creek, have you? ▷ takes floor from S4

79.5–89.7 **S2** [turn] It was about four years ago. Took three of us to get the thing loose—it was so caked in mud, we joked it looked like a modern sculpture. People who walked by started applauding as we rolled it up the embankment. ▷ yields to S1

89.6–92.0 **S1** [turn] Wow, uh, that must've felt pretty epic. ▷ yields to S3

92.4–100.8 **S3** [turn] Yeah, the crowd energy is always surprising, right? I had a little girl hand me dandelions as a thank you once. It made the whole day. ▷ yields to S2

94.7–95.1 **S4** [backchannel] Awww

98.2–99.2 **S1** [backchannel] That's sweet.

100.9–105.1 **S2** [turn] Yeah, you get these moments that stick with you, y'know? Reminds you why you keep showing up. ▷ yields to S3

105.3–115.6 **S3** [turn] Absolutely. Okay, so, switching gears, is there anything about this upcoming river event you're, uh, curious or worried about? Any first-timers need tips? ▷ yields to S1

115.6–124.9 **S1** [turn] Well, I, um, I haven't led a group before, so I'd really appreciate, like, any advice on keeping people motivated if the weather turns bad or, uh, if we get a low turnout. ▷ holds, pauses

118.8–120.3 **S4** [failed claim] Oh, I got you! But first—

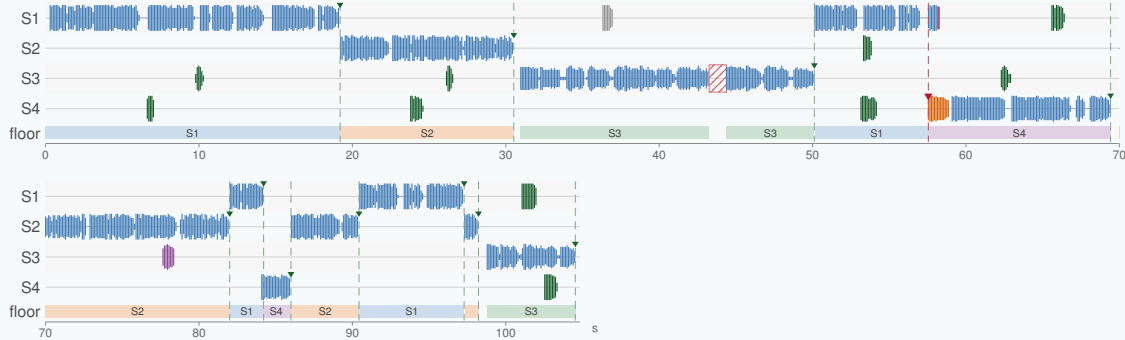
125.6–129.6 **S1** [turn] —and, uh, just, if anybody has a go-to pep talk or something, that'd be awesome. ▷ yields to S4

129.5–142.9 **S4** [turn] I was gonna say, snacks are like, ninety percent of morale. The other ten is music. Seriously, you play '80s hits or bring cookies, everyone's suddenly an eager beaver. I keep a stash of silly hats for emergencies. ▷ yields to S3

136.9–138.4 **S2** [laughter] Haha, silly hats.

142.9–149.1 **S3** [turn] Perfect, I think that's our official team strategy: hats, cookies, and, uh, stubborn optimism. ▷ yields, end

test split · tone: thoughtful and contemplative · noise: studio room tone, 28.0 dB SNR · voices: S1 M, S2 M, S3 M, S4 M
 EOT 11 yields / 1 holds · INT 1 interruptions / 11 negatives (10 backchannels, 1 laughter) · 1 failed claims (excluded)



Transcript of test_00107

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–19.2 **S1** [turn] I was thinking about the setup for the food drive last year, you know, how we all got there so early and it was still sort of foggy out. I remember standing by those tables trying to tape the sign up, and it felt like, even before people arrived, there was this weird, good energy. Like we all knew, okay, today's gonna be busy, but it's, uh, for the right reasons. ▷ yields to S2

6.6–7.1 **S4** [backchannel] Right

9.8–10.3 **S3** [backchannel] Yeah

19.2–30.5 **S2** [turn] I, uh, I always get nervous before those things. I remember you had that box of gloves, and for some reason just handing those out, it— kind of made things feel official, like we all had a part. ▷ yields to S3

23.8–24.6 **S4** [backchannel] Absolutely

26.1–26.6 **S3** [backchannel] Mm-hm

31.0–43.3 **S3** [turn] I remember Elaine showing up right in the middle of that rush. She'd, uh, brought her guitar—I guess thinking there'd be downtime. She just started singing to the line and suddenly everyone, I don't know, loosened up? Even the folks waiting looked less anxious. ▷ holds, pauses

36.3–37.0 **S1** [failed claim] Oh yeah—

44.4–50.1 **S3** [turn] —She played that old Bob Dylan song, I think. The one about shelter from the storm. It felt kinda perfect. ▷ yields to S1

50.2–58.2 **S1** [turn] Yeah! And there was that little kid in the red jacket, remember? He tried strumming along, on, um, a plastic bat he had. I just remember thinking, you don't plan for those bits, but they're the ones that stick with you. ▷ cut off

53.2–54.2 **S4** [backchannel] Right, that's true

53.3–53.9 **S2** [backchannel] Totally

57.5–58.9 **S4** [interruption] What I remember most— ▷ takes floor from S1

59.1–69.4 **S4** [turn] —was after, when we were packing up and, uh, nobody really wanted to leave yet. It's like, those moments cleaning up, just talking, you realize how rare it is to feel that... connected. Even after the crowd's gone. ▷ yields to S2

62.3–63.0 **S3** [backchannel] Yeah.

65.6–66.5 **S1** [backchannel] Mm, exactly

70.0–82.0 **S2** [turn] I think... like, for me, that's almost the best part? I mean, doing the event, that's great, but afterward it's, uh, all the little stories from the day coming out. Like when Sarah tripped over that box but just kept laughing— ▷ yields to S1

77.7–78.4 **S3** [laughter] Haha yeah

82.0–84.2 **S1** [turn] Yeah, she just kept going! Nothing phased her. ▷ yields to S4

84.1–86.0 **S4** [turn] That's real resilience, right there. ▷ yields to S2

86.0–90.4 **S2** [turn] I wish I could be that relaxed when something awkward happens. I always freeze up— ▷ yields to S1

90.5–97.3 **S1** [turn] No, honestly, you keep things moving in your own way. Like when the truck was late? You just kept sorting cans like nothing was wrong. ▷ yields to S2

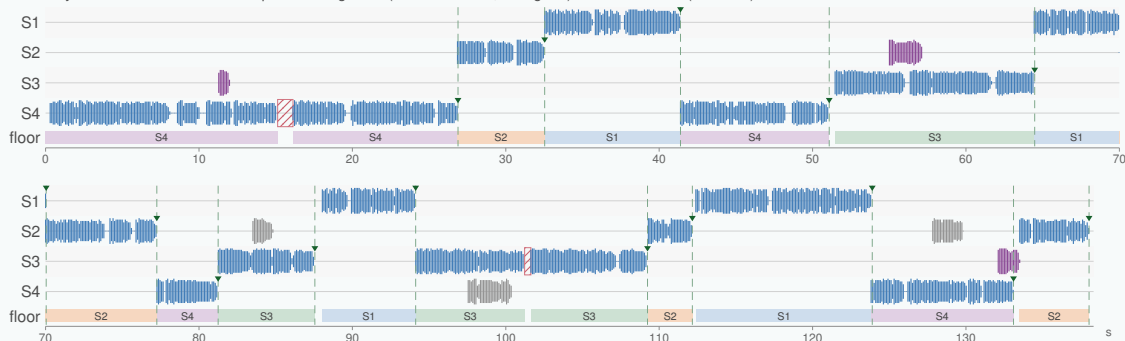
97.3–98.2 **S2** [turn] Oh, um—thanks. ▷ yields to S3

98.8–104.5 **S3** [turn] That's how it goes. Everyone's got their bit. Some do the singing, some just keep the cans rolling. All of it matters. ▷ yields, end

101.1–102.1 **S1** [backchannel] Couldn't agree more.

102.5–103.4 **S4** [backchannel] Exactly.

test split · tone: enthusiastic and energetic · noise: outdoor park, 20.3 dB SNR; distant speech · voices: S1 M, S2 F, S3 M, S4 F
 EOT 15 yields / 2 holds · INT 0 interruptions / 3 negatives (0 backchannels, 3 laughter) · 3 failed claims (excluded)



Transcript of test_00116

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–15.1 **S4** [turn] Okay, before we start on the practice problems, I have to tell you guys about what happened last night when I tried to memorize all those psychology terms. It was actually hilarious. So I'm sitting at my desk, right, and my cat suddenly jumps onto my notes and literally— ▷ holds, pauses

11.3–12.0 **S3** [laughter] Haha no way—

16.1–26.9 **S4** [turn] —starts pawing at the flashcards like they're her prey or something. Next thing I know, half my cards are under the dresser, and I'm crawling on the floor just to get to, like, 'operant conditioning.' > yields to S2

26.9–32.5 **S2** [turn] Oh my gosh, that sounds like my dog. He ate my chemistry notes once. I had to go back and rewrite everything! > yields to S1

32.5–41.4 **S1** [turn] That—um, that actually sounds kind of stressful, but it's, uh, kinda funny too. Did you remember any of the terms after that, or did it just make it worse? > yields to S4

41.4–51.1 **S4** [turn] Honestly, I think crawling around on the floor probably just drilled 'stimulus generalization' into my brain more than anything. Like, now I associate it with dust bunnies! > yields to S3

51.5–64.5 **S3** [turn] Classic, I wish I had a good animal story, but my goldfish literally does nothing. Though last year, I actually fell asleep on my textbook and woke up with highlighter lines on my cheek. I think that counts as osmosis learning, right? > yields to S1

55.0–57.1 **S2** [laughter] Haha, sure, if only it worked that way.

64.4–70.0 **S1** [turn] Yeah, if I could just, uh, hold my notebook and absorb it all, exams would be so much less stressful. > yields to S2

70.0–77.2 **S2** [turn] Honestly though, speaking of stress, I always get so nervous before big exams. Like, my hands get all shaky. Does that happen to anyone else? > yields to S4

77.3–81.2 **S4** [turn] Totally! I get so much adrenaline I end up bouncing my leg the whole time. > yields to S3

81.3–87.6 **S3** [turn] Yeah, me too, but I kind of lean into it. Like, I just tell myself, 'Alright, bring it on,' even if I don't feel ready. > yields to S1

83.5–84.8 **S2** [failed claim] That's a good attitude.

88.0–94.1 **S1** [turn] I kind of freeze up, honestly. Like, I always blank on the first question and have to, um, just skip ahead. > yields to S3

94.1–101.2 **S3** [turn] Yeah, that's real. I used to do that too, but now I just write anything for the first one, even if it's nonsense, just to get moving. > holds, pauses

97.5–100.4 **S4** [failed claim] Wait, you just—so you literally write random stuff at first?

101.6–109.2 **S3** [turn] —well, I mean, not, like, total garbage, but I write something related, just to stop staring at the page. It kind of unclogs my brain. > yields to S2

109.3–112.2 **S2** [turn] That actually makes sense. Maybe I'll try that on Friday. > yields to S1

112.4–123.9 **S1** [turn] Wait, but, uh, has anyone ever actually run out of time because of freezing up? Because last semester, during our stats final, I spent forever on the first problem and then totally panicked for the rest. > yields to S4

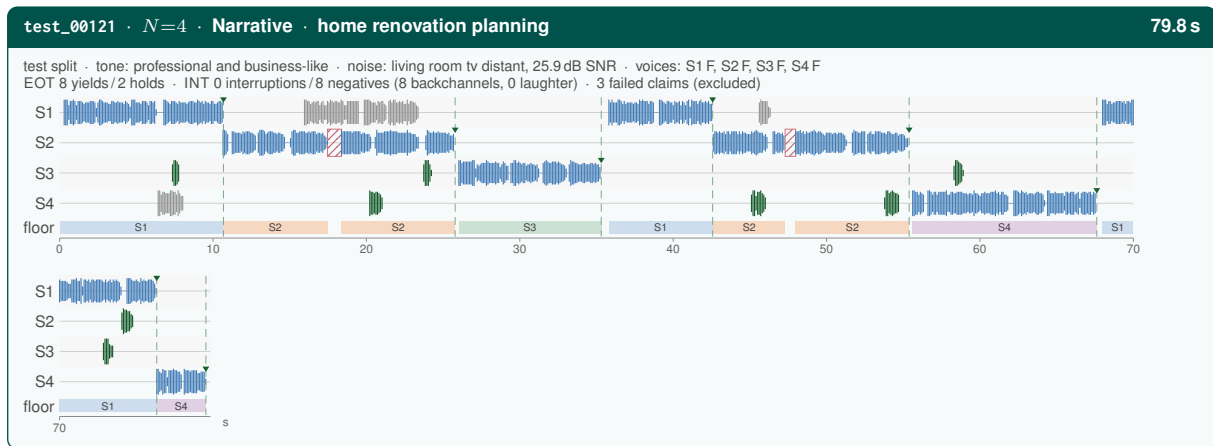
123.8–130.1 **S4** [turn] Oh, totally, I did that with biology! I got stuck on that weird cell diagram and didn't even finish the essays.

127.8–129.8 **S2** [failed claim] Yeah, that cell diagram was brutal.

130.2–133.1 **S4** [turn] Right? I was like, 'What even is mitochondria doing here?' > yields to S2

132.1–133.5 **S3** [laughter] Haha, powerhouses everywhere.

133.5–138.0 **S2** [turn] You know what, though? All of this just makes me feel a little less alone about freaking out before exams. > yields, end



Transcript of test_00121

start—end (s) speaker [ground-truth label] text > outcome

0.3–10.7 **S1** [turn] Alright, so the last time I did a kitchen reno, the electrician showed up a full week late, and everything snowballed from there. If we're going to take out this wall and redo the lighting, I want a tighter schedule than that. > yields to S2

6.4–8.1 **S4** [failed claim] Absolutely, that makes sense.

7.4–7.8 **S3** [backchannel] mm-hm

10.7–17.5 **S2** [turn] Right, but if we sequence the demo and the electrical rough-in together, we can actually reduce the idle time by— > holds, pauses

16.0–23.5 **S1** [failed claim] —But that only works if the permits clear right away, which last spring, they did not. Remember my neighbor's project? The inspector just didn't show for days.

18.4–25.8 **S2** [turn] —as long as we file electronically, the city says they're averaging forty-eight hours now. At least that's what their portal says. > yields to S3

20.2–21.1 **S4** [backchannel] Oh really

23.7–24.3 **S3** [backchannel] Yeah

26.0–35.3 **S3** [turn] Last time I did the bathroom, the inspector rescheduled twice, but I got a text both times at least. What about supply delays, though? Cabinet lead times are insane right now. > yields to S1

35.8–42.6 **S1** [turn] Exactly, and that's where people lose weeks they never planned for. I waited twelve weeks on a sink once, and the plumber wouldn't finish without it. > yields to S2

42.6–47.3 **S2** [turn] So maybe we should lock in all the custom orders before anything else, just to be safe. > holds, pauses

45.1–46.1 **S4** [backchannel] That's smart.

45.6–46.3 **S1** [failed claim] Yeah, but—

48.0–55.4 **S2** [turn] —and I have a supplier who got my last client's counters in a month, so maybe we're not totally at the mercy of global shipping this time. > yields to S4

53.8–54.8 **S4** [backchannel] Good to know.

55.6–67.6 **S4** [turn] And for what it's worth, every single client I've had who scheduled backup days for inspections or deliveries ended up less stressed. I always recommend building some buffer in—that way, there's room if, you know, someone gets sick or the weather acts up. > yields to S1

58.3–58.9 **S3** [backchannel] Mhm.

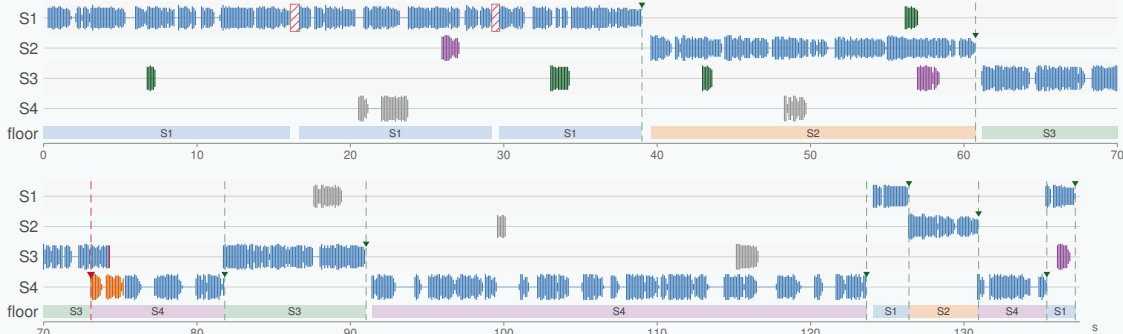
68.0–76.3 **S1** [turn] Alright, so let's put together a timeline with buffer days, order everything up front, and flag the permit submission as day one. If anyone flakes, I want it in writing. > yields to S4

72.9–73.5 **S3** [backchannel] Yep.

74.1–74.8 **S2** [backchannel] Sounds good.

76.3–79.5 **S4** [turn] That's a solid plan, everyone. I think this will go much smoother. > yields, end

test split · tone: professional and business-like · noise: school hallway, 26.7 dB SNR; chairs scraping, laughter offmic · voices: S1 F, S2 M, S3 F, S4 F
EOT 9 yields/2 holds · INT 1 interruptions/7 negatives (4 backchannels, 3 laughter) · 5 failed claims (excluded)



Transcript of test_00136

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–16.1 **S1** [turn] Thank you all for having me today. So, um, you asked about a challenge I overcame at work. Last year, right before our big product launch, our lead designer quit—just vanished! So I had to basically jump in and, well, become the unofficial fire chief for a week. ▷ holds, pauses

6.8–7.3 **S3** [backchannel] Oh wow.

16.7–29.2 **S1** [turn] Yeah, it was—let's say, a crash course in crisis management. But uh, the fun part was rallying the team. I started each day with these really bad motivational quotes, like, 'Coffee first, panic later.' ▷ holds, pauses

20.5–23.8 **S4** [failed claim] Did it work? Or did everyone quit by day two?

26.0–27.1 **S2** [laughter] Haha, that's great.

29.7–39.0 **S1** [turn] —Surprisingly, no one quit! We actually hit our deadline, and, uh, I learned I can survive on four hours of sleep if sufficiently bribed with pastries. ▷ yields to S2

33.1–34.3 **S3** [backchannel] Absolutely relatable.

39.6–44.2 **S2** [turn] That's a pretty heroic week. Uh, I actually had something similar happen a few years ago—

43.0–43.6 **S3** [backchannel] Oh really?

44.4–54.9 **S2** [turn] Our warehouse system crashed two days before Black Friday. I stayed up basically all night on the phone with IT, trying to fix things. At some point, I started hallucinating error codes in my dreams.

48.3–49.8 **S4** [failed claim] Nightmare fuel for sure.

55.0–60.8 **S2** [turn] Yeah! But, uh, when it was all over, my manager sent me a mug that said 'Error 404: Sleep Not Found.' ▷ yields to S3

56.2–57.1 **S1** [backchannel] That's so good!

57.0–58.4 **S3** [laughter] Haha, I need a mug like that.

61.2–74.3 **S3** [turn] So, for me—uh, my story's more on the sales side. There was this one campaign, honestly, it bombed at first. We weren't getting any leads, and my boss was, um, not thrilled. I just started calling people to figure out what we were missing. ▷ cut off

73.1–75.3 **S4** [interruption] Did anyone actually pick up? ▷ takes floor from S3

75.3–81.8 **S4** [turn] Because I swear, if I see an unknown number, I just let it ring out of spite. ▷ yields to S3

81.7–91.0 **S3** [turn] Oh, same! But, uh, I actually got through to a few, and it turned out we'd missed a key feature in our campaign. Once we fixed that, suddenly leads started rolling in. ▷ yields to S4

87.6–89.5 **S1** [failed claim] Nice, that's persistence right there.

91.4–110.1 **S4** [turn] Well, I guess that leaves me. Uh, my best workplace story is probably less heroic than yours. Once, my team had to present to the board, and our projector just, you know, erupted into smoke mid-slide.

99.6–100.2 **S2** [failed claim] Oh no—

110.3–123.6 **S4** [turn] —So, I just handed out the print-outs and told them, 'At least these can't catch fire.' Not my slickest save, but hey, everyone remembered that pitch. ▷ yields to S1

115.2–116.7 **S3** [failed claim] Memorable's what matters, right?

124.1–126.4 **S1** [turn] Honestly, that's the kind of story boards love. ▷ yields to S2

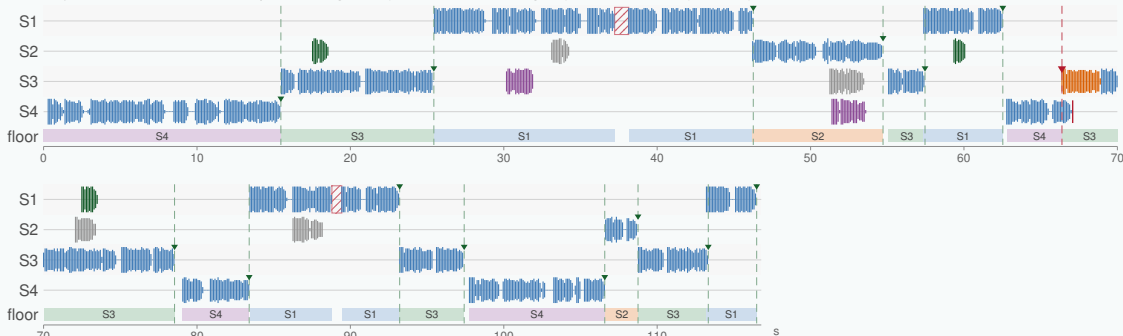
126.4–131.0 **S2** [turn] Yeah, if you can survive a literal smoke bomb in the boardroom, I think you're good under pressure. ▷ yields to S4

130.9–135.4 **S4** [turn] Thanks. If your projector spontaneously combusts, call me. ▷ yields to S1

135.3–137.3 **S1** [turn] Deal. Only if you bring the pastries. ▷ yields, end

136.1–136.9 **S3** [laughter] Haha, agreed!

test split · tone: enthusiastic and energetic · noise: studio room tone, 29.9 dB SNR; keyboard clatter · voices: S1 M, S2 F, S3 F, S4 M
EOT 14 yields/2 holds · INT 1 interruptions/5 negatives (3 backchannels, 2 laughter) · 4 failed claims (excluded)



Transcript of test_00145

start-end (s) speaker [ground-truth label] text ▷ outcome

03.1–15.5 **S4** [turn] Alright, everyone, thanks for making time. I know things have gotten kind of wild with all our schedules, so I thought we could check in. Roommate summit. Let's kick off with stories—what's something that's made life easier or harder around here lately? **> yields to S3**

15.5–25.4 **S3** [turn] Oh, I love this! So, um, actually having that shoe rack by the door has made a huge difference for me. Like, I'm not tripping over sneakers every morning anymore, and it honestly feels like tiny magic. **> yields to S1**

17.5–18.6 **S2** [backchannel] Yeah, same.

25.5–37.2 **S1** [turn] Totally. I actually keep losing my own shoes less since we did that. Uh, but what really helped me last week was when someone—uh, I think it was you. Maya—cleaned out the fridge. I found my leftovers for once. **> holds, pauses**

30.2–31.9 **S3** [laughter] Guilty! But in a good way, I hope.

33.1–34.3 **S2** [failed claim] So, about that—

38.1–46.3 **S1** [turn] —no, it was honestly a game changer. I know I keep talking about being forgetful, but like, half my problems are just not seeing things. So, thank you. **> yields to S2**

46.2–54.7 **S2** [turn] Okay, uh, I was trying to say—can we maybe, like, make that a every-two-weeks thing? Because stuff keeps appearing in the back and I swear the salsa is growing a beard in there. **> yields to S3**

51.3–53.5 **S3** [failed claim] Oh my gosh, yes! I saw that jar too—

51.4–53.6 **S4** [laughter] Haha, I've been scared to touch it for a week.

55.1–57.5 **S3** [turn] I vote yes, fridge cleaning on the calendar. **> yields to S1**

57.4–62.5 **S1** [turn] Agreed. I can help, but someone else needs to check expiration dates. I'm, uh, not brave enough. **> yields to S4**

59.3–60.1 **S2** [backchannel] Can do.

62.8–67.1 **S4** [turn] Alright, two-week fridge check, I'll put it on the roster. Speaking of schedules— **> cut off**

66.4–68.9 **S3** [interruption] Wait, before you get into that, I just wanted to share— **> takes floor from S4**

69.0–78.5 **S3** [turn] I was talking to my old college roommate and, um, she said her house never argues because they do group dinners once a week. And it got me thinking, like, we could maybe try it? I know it sounds cheesy but— **> yields to S4**

72.1–73.5 **S2** [failed claim] Actually, I like it.

72.5–73.5 **S1** [backchannel] That sounds nice.

79.1–83.4 **S4** [turn] Yeah, I'd be down for that. Should we schedule it or see who's around each week? **> yields to S1**

83.5–88.8 **S1** [turn] If we put it on the calendar, I'll actually remember. And maybe alternate who cooks, so no one gets stuck? **> holds, pauses**

86.3–88.3 **S2** [failed claim] Wait—do we rotate, or teams?

89.4–93.2 **S1** [turn] —I like teams, actually. Pairing up would be fun and less overwhelming. **> yields to S3**

93.2–97.4 **S3** [turn] Yesss, and then we can try new recipes too. I'll even make dessert if we're on a team. **> yields to S4**

97.8–106.6 **S4** [turn] Alright, group dinners on the calendar, paired teams, and biweekly fridge cleanse. Anything else before I, uh, officially wrap our summit? **> yields to S2**

106.6–108.8 **S2** [turn] No, that covers it. Thanks, all. **> yields to S3**

108.8–113.3 **S3** [turn] Honestly, this was super helpful. I feel, like, weirdly excited to clean the fridge now. **> yields to S1**

113.2–116.5 **S1** [turn] Haha, I'll believe that when I see it. But, yeah, thanks everyone. **> yields, end**

Transcript of test_00159

0.3–19.8 **S1** [turn] Alright, so we're still trying to decide where to go tonight, yeah? I know we tossed around that new Thai place, but honestly, the last time I tried something super spicy on a group night, it kind of, uh, backfired. I mean, not to be dramatic, but I ended up spending most of the meal just asking for more water and apologizing to the waitress for, like, sweating through my shirt. [▷ yields to S2](#)

12.4–14.0 **S3** [failed claim] Oh my god, that's so relatable!

20.2–25.1 **S2** [turn] I-I mean, Thai is good but, um, I kind of have a story too. The last time I went out for sushi—

21.7–23.1 **S4** [backchannel] Ooooh, sushi disaster?

25.3–38.7 **S2** [turn] —sort of! I accidentally ate like a whole chunk of wasabi because I thought it was avocado. I-uh, I didn't want to be rude so I just quietly panicked for like three minutes, tears running down my face. The chef saw me and brought me milk, which actually... made it worse. [▷ yields to S3](#)

29.3–31.5 **S3** [failed claim] Oh no, not the wasabi move! That's brutal.

29.6–31.0 **S4** [laughter] Haha, the silent suffering...

39.2–56.6 **S3** [turn] Okay, but you know what I think would be fun? Tapas. Hear me out—there was this one time my friends and I just, like, ordered the whole menu at this tiny place. Everybody was just reaching around, grabbing stuff, and by the end we had no idea who actually paid for what. I think the waiter just gave up and split it four ways.

44.9–46.5 **S1** [failed claim] Yeah, I like the idea of sharing—

56.7–63.8 **S3** [turn] —right? It's just so social. And you get to taste everything instead of committing to one thing and then getting, like, major food envy. [▷ yields to S1](#)

63.9–74.9 **S1** [turn] —but I also remember once at a tapas place, no one touched the octopus, and it just sat there. We were all too polite to say we didn't want it, so it became this centerpiece of uneaten food until the end. [▷ yields to S2](#)

68.9–70.3 **S4** [laughter] Haha, octopus stand-off!

75.3–79.6 **S2** [turn] I'd be that person pretending to like it but secretly pushing it around my plate. [▷ yields to S4](#)

77.0–79.5 **S3** [failed claim] Same! I'm not brave enough for fancy sea creatures.

79.8–89.1 **S4** [turn] You guys! You have to try new things at least once, though. Like, last year, I ended up at this Ethiopian place for a birthday and I still crave their spicy lentils. Also, eating with your hands? Very freeing. But also, um, incredibly messy for me. My shirt did not survive—I probably looked like a toddler. [▷ cut off](#)

85.9–88.1 **S3** [laughter] Haha, actually, Ethiopian sounds amazing!

88.5–90.8 **S2** [interruption] Wait, I've never even had Ethiopian— [▷ takes floor from S4](#)

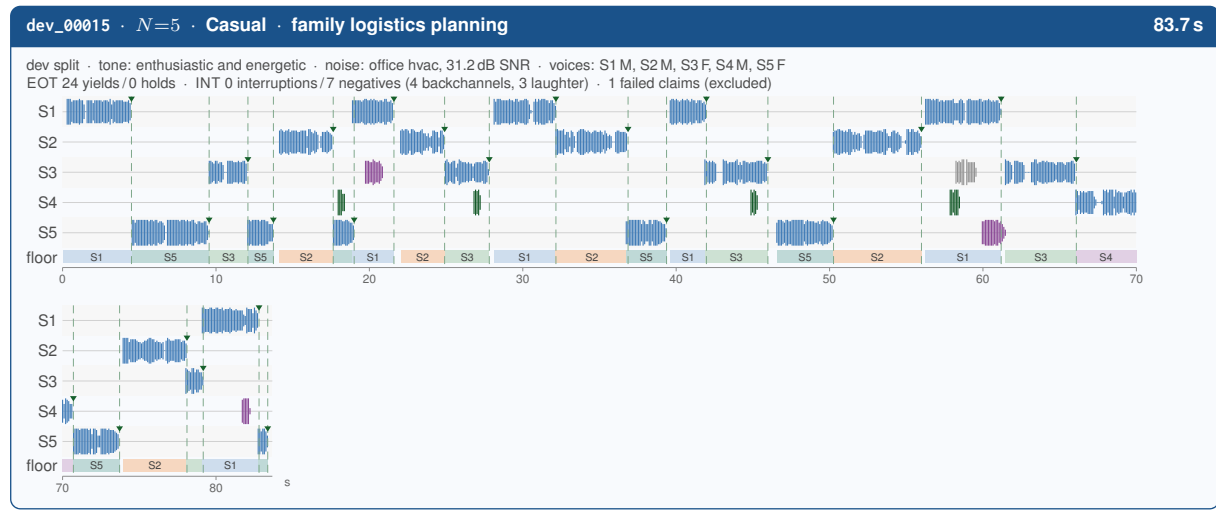
90.8–98.9 **S2** [turn] —is it, like, super spicy or more, um, savory? I just want to be prepared if we're going full adventure tonight. And does everyone actually eat with their hands? $\triangleright$ yields to S4

98.9–108.8 **S4** [turn] Oh yeah! It's all about the injera bread. You scoop everything up with it. Definitely messy, but in a good way. I promise, even if you're awkward at first, it's part of the charm. $\triangleright$ yields to S1

109.0–116.9 **S1** [turn] That honestly sounds great. I'm in if you all are. But I'm claiming a seat farthest from the spicy stuff—just in case history repeats itself. $\triangleright$ yields, end

114.5–115.0 **S3** [backchannel] Deal!

K.4 Conversations with $N = 5$ speakers



Transcript of dev_00015

start—end (s) speaker [ground-truth label] text $\triangleright$ outcome

0.3–4.5 **S1** [turn] Alright, so just checking, are we all clear on who's doing the airport run Saturday morning? $\triangleright$ yields to S5

4.5–9.6 **S5** [turn] That's me, I volunteered last night. I've already booked the Zipcar for seven, so we're set. $\triangleright$ yields to S3

9.6–12.1 **S3** [turn] Perfect, thank you. And you've got the address, right? $\triangleright$ yields to S5

12.1–13.7 **S5** [turn] Yeah, I saved it on Google Maps. $\triangleright$ yields to S2

14.1–17.6 **S2** [turn] Wait, is that the terminal by the new parking lot, or the old one? $\triangleright$ yields to S5

17.7–19.0 **S5** [turn] Uh, the new one. Terminal C. $\triangleright$ yields to S1

18.0–18.4 **S4** [backchannel] mm-hm

18.9–21.6 **S1** [turn] Cool, cool, just making sure no one's getting lost again. $\triangleright$ yields to S2

19.7–20.9 **S3** [laughter] Ha! Like last time—

22.0–24.9 **S2** [turn] Hey, I was following your directions. Not my fault. $\triangleright$ yields to S3

25.0–27.8 **S3** [turn] Well, this time, maps are being used, so we should be fine. $\triangleright$ yields to S1

26.8–27.3 **S4** [backchannel] yeah

28.1–32.2 **S1** [turn] Okay, next thing, who's bringing snacks for the drive up? I still have that big cooler. $\triangleright$ yields to S2

32.2–36.9 **S2** [turn] I'll do chips and fruit, but someone else has to remember the ice packs. I always forget. $\triangleright$ yields to S5

36.7–39.4 **S5** [turn] I can handle the ice packs. I'll put it on my list. $\triangleright$ yields to S1

39.6–42.0 **S1** [turn] Great, thanks. Uh, what about the car seat for Max? $\triangleright$ yields to S3

41.9–46.0 **S3** [turn] I'm bringing it. I'll double-check it's strapped in properly before you guys get here. $\triangleright$ yields to S5

44.9–45.4 **S4** [backchannel] uh-huh

46.6–50.3 **S5** [turn] So, what time is everyone getting here in the morning, like, realistically? $\triangleright$ yields to S2

50.2–56.0 **S2** [turn] If we want to leave by seven-fifteen, people need to be here by seven, seriously. Not seven-twelve. $\triangleright$ yields to S1

56.3–58.9 **S1** [turn] I promise I'll actually be early for once. Swear.

57.9–58.5 **S4** [backchannel] no way

58.2–59.6 **S3** [failed claim] You say that every year—

59.0–61.2 **S1** [turn] —no, really! I'll set three alarms this time. $\triangleright$ yields to S3

60.0–61.5 **S5** [laughter] Haha, I'll believe it when I see it.

61.5–66.1 **S3** [turn] Alright, moving on! Who's actually got the reservation confirmation for the cabin? $\triangleright$ yields to S4

66.1–70.7 **S4** [turn] I printed the email. I can send a photo if people want it on their phones. $\triangleright$ yields to S5

70.7–73.7 **S5** [turn] Can you just drop it in the group chat? That way we all have it. $\triangleright$ yields to S2

74.0–78.1 **S2** [turn] Alright! So basically, if nobody oversleeps, nothing can go wrong, right? $\triangleright$ yields to S3

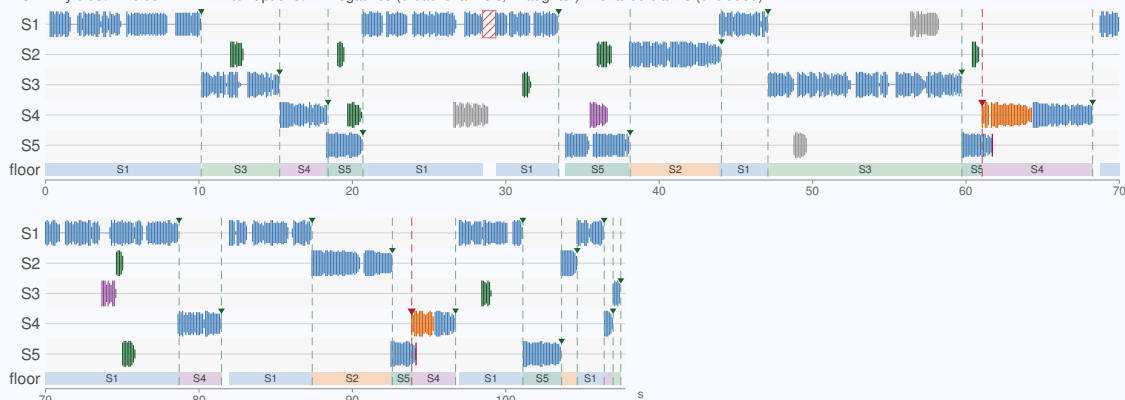
78.0–79.2 **S3** [turn] Famous last words. $\triangleright$ yields to S1

79.1–82.8 **S1** [turn] Knock on wood! But yeah, we're actually organized for once. Group high five? $\triangleright$ yields to S5

81.7–82.2 **S4** [laughter] haha

82.8–83.4 **S5** [turn] Go team. $\triangleright$ yields, end

dev split · tone: casual and friendly · noise: outdoor park, 31.0 dB SNR; chairs scraping · voices: S1 F, S2 F, S3 M, S4 F, S5 M
EOT 21 yields / 1 holds · INT 2 interruptions / 11 negatives (9 backchannels, 2 laughter) · 3 failed claims (excluded)

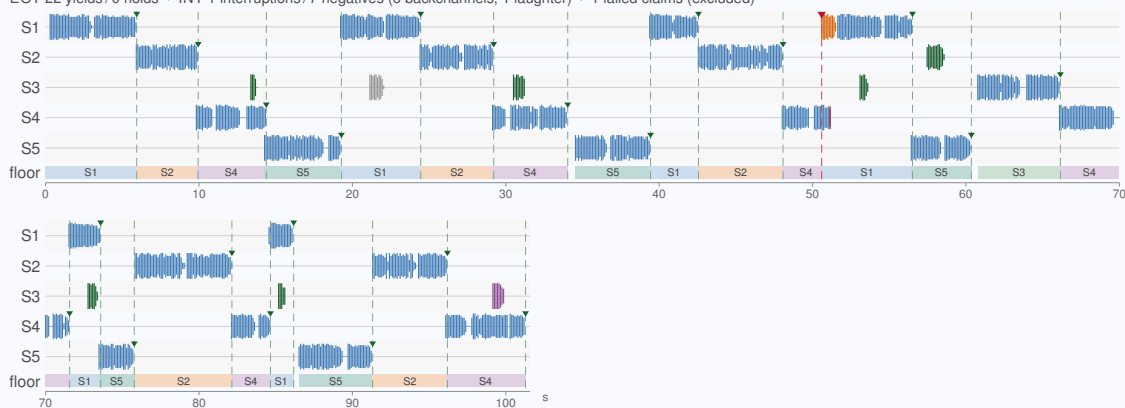


Transcript of dev_00045

start—end (s) speaker [ground-truth label] text > outcome

0.3–10.2 **S1** [turn] Alright, let's get started. So, the practice exam says, uh, question three: explain the difference between diffusion and osmosis. Who remembers the main point there? > yields to S3
10.2–15.3 **S3** [turn] Osmosis is just, like, water, right? Everything else is just called diffusion. > yields to S4
12.1–12.9 **S2** [backchannel] Yeah exactly.
15.3–18.4 **S4** [turn] But isn't it also through a membrane? That's the osmosis part. > yields to S5
18.3–20.7 **S5** [turn] Right, like the cell membrane is semi-permeable— > yields to S1
19.0–19.5 **S2** [backchannel] Mm-hmm.
19.7–20.6 **S4** [backchannel] Exactly.
20.6–28.5 **S1** [turn] So if you're writing the answer, don't just say water moves, make sure you put that it's across a semi-permeable membrane. Otherwise you'll lose— > holds, pauses
26.6–28.9 **S4** [failed claim] Wait, so just water? Or like any solvent?
29.4–33.5 **S1** [turn] —lose the point, yeah, just water in bio, unless they specify otherwise. > yields to S5
31.1–31.6 **S3** [backchannel] Cool.
33.9–38.1 **S5** [turn] Haha, I always mess that up. I write like this whole paragraph and forget that part. > yields to S2
35.5–36.7 **S4** [laughter] Haha, me too.
36.0–37.0 **S2** [backchannel] Yeah, same honestly.
38.1–44.1 **S2** [turn] Okay, wait, can we talk through question five? The one about enzyme inhibition. I always get confused. > yields to S1
43.9–47.1 **S1** [turn] Sure, so you want competitive vs non-competitive, right? > yields to S3
47.1–50.4 **S3** [turn] One is like musical chairs and one is sabotage.
48.8–49.6 **S5** [failed claim] Wait can you—
50.5–59.7 **S3** [turn] —Competitive is musical chairs. Inhibitor competes for the active site. Non-competitive is sabotage, blocks somewhere else and changes the shape. > yields to S5
56.4–58.3 **S1** [failed claim] Yeah, that's actually a good analogy.
59.7–61.7 **S5** [turn] Haha, okay, that's actually helpful. > cut off
60.4–60.9 **S2** [backchannel] Right.
61.1–64.3 **S4** [interruption] But wait, does non-competitive inhibition lower Vmax? > takes floor from S5
64.4–68.3 **S4** [turn] I always mix up which one changes Km and which one changes Vmax. > yields to S1
68.7–78.7 **S1** [turn] So, non-competitive lowers Vmax, because it takes some of the enzymes out of commission. Competitive just increases Km, since you need more substrate to outcompete. > yields to S4
73.7–74.6 **S3** [laughter] Ha, makes sense.
74.6–75.1 **S2** [backchannel] Mm-hmm.
75.0–75.9 **S5** [backchannel] Yeah, got it.
78.6–81.5 **S4** [turn] Perfect, okay, I think I can finally remember that. > yields to S1
82.0–87.4 **S1** [turn] Alright, last one—photosynthesis versus cellular respiration, anyone wanna jump in? > yields to S2
87.4–92.6 **S2** [turn] Photosynthesis stores energy, respiration releases it. Sunlight in, ATP out. > yields to S5
92.5–94.2 **S5** [turn] Yeah, they're basically opposites. > cut off
93.9–95.3 **S4** [interruption] Except plants do both. > takes floor from S5
95.4–96.7 **S4** [turn] I always forget that. > yields to S1
97.0–101.1 **S1** [turn] Yeah, exactly, so don't put 'plants only do photosynthesis', remember both. > yields to S5
98.4–99.1 **S3** [backchannel] Gotcha.
101.1–103.6 **S5** [turn] Alright, I feel a little better about this exam now. > yields to S2
103.6–104.7 **S2** [turn] Yeah, thanks, everyone. > yields to S1
104.6–106.4 **S1** [turn] You all did great, see you Thursday! > yields to S4
106.4–107.0 **S4** [turn] Bye! > yields to S3
107.0–107.5 **S3** [turn] Later. > yields, end

test split · tone: curious and inquisitive · noise: outdoor park, 22.6 dB SNR · voices: S1 M, S2 F, S3 M, S4 F, S5 M
 EOT 22 yields / 0 holds · INT 1 interruptions / 7 negatives (6 backchannels, 1 laughter) · 1 failed claims (excluded)



Transcript of test_00009

start-end (s) speaker [ground-truth label] text > outcome

0.3–5.9 **S1** [turn] Alright so, did anyone catch the Knicks game last night? Man, I honestly can't believe that buzzer beater. > yields to S2
 5.9–9.9 **S2** [turn] Yeah, it was wild! I actually thought they'd messed it up with that last turnover. > yields to S4
 9.8–14.4 **S4** [turn] Oh my gosh, same! I was literally yelling at the TV. My cat looked terrified. > yields to S5
 13.4–13.8 **S3** [backchannel] mm-hm
 14.3–19.3 **S5** [turn] Honestly, I thought the refs were gonna call an offensive foul on that last drive, but they let it go. > yields to S1
 19.2–22.1 **S1** [turn] Do you think it was the right call though? That push looked a little—
 21.1–22.1 **S3** [failed claim] Wait, actually—
 22.2–24.5 **S1** [turn] —iffy from the replay angle, if I'm being honest. > yields to S2
 24.4–29.2 **S2** [turn] Right? I kept pausing and rewatching that slo-mo. It looked like there was some contact, but— > yields to S4
 29.1–34.0 **S4** [turn] Haha, same! I love going frame by frame, like it makes me some expert or something. > yields to S5
 30.5–31.3 **S3** [backchannel] exactly
 34.5–39.4 **S5** [turn] I think, like, in real speed, it looked fine. If they called every little push, there'd be no flow. > yields to S1
 39.4–42.6 **S1** [turn] Yeah but sometimes they're just, like, super inconsistent, you know? > yields to S2
 42.6–48.1 **S2** [turn] For sure, you never know what you're gonna get. Oh, and can we talk about Hart's defense? He was everywhere. > yields to S4
 48.0–51.1 **S4** [turn] Right! He totally shut down the paint. Like, the one block he had— > cut off
 50.6–51.6 **S1** [interruption] Oh, that was insane! > takes floor from S4
 51.6–56.5 **S1** [turn] He literally came out of nowhere, like a ghost or something. I rewind that like three times. > yields to S5
 53.1–53.7 **S3** [backchannel] uh-huh
 56.4–60.4 **S5** [turn] And he hustled after every loose ball. I love seeing that kind of energy— > yields to S3
 57.5–58.6 **S2** [backchannel] Yeah, totally!
 60.8–66.2 **S3** [turn] Actually, I kinda thought he might foul out at one point. He was really aggressive on D there in the third. > yields to S4
 66.1–71.6 **S4** [turn] Wait wait, but did y'all see that guy in the crowd with the giant foam hat? Like, what even was that thing? > yields to S1
 71.5–73.6 **S1** [turn] Haha, yeah! I thought he was some mascot at first. > yields to S5
 72.8–73.4 **S3** [backchannel] no way
 73.5–75.8 **S5** [turn] I totally missed that, was it on the broadcast? > yields to S2
 75.8–82.1 **S2** [turn] Just for, like, a second during the timeout, but he was dancing! My friend texted me a screenshot, it was hilarious. > yields to S4
 82.1–84.7 **S4** [turn] Haha, people get so creative. Props to him. > yields to S1
 84.6–86.2 **S1** [turn] You gotta love New York fans, man. > yields to S5
 85.2–85.7 **S3** [backchannel] right
 86.5–91.3 **S5** [turn] So, uh, do you all think they've got a shot at the playoffs this year? Or is it still too early to say? > yields to S2
 91.3–96.2 **S2** [turn] I mean, if they keep up this kind of effort, I'm hopeful. But you never know with their injuries. > yields to S4
 96.1–101.3 **S4** [turn] Yeah, but I wanna believe! Plus, if Brunson keeps playing like that, anything's possible. > yields, end
 99.2–99.9 **S3** [laughter] haha right

test split · tone: enthusiastic and energetic · noise: studio room tone, 26.5 dB SNR; chairs scraping, paper shuffle · voices: S1 M, S2 F, S3 M, S4 F, S5 F
 EOT 22 yields / 1 holds · INT 2 interruptions / 6 negatives (6 backchannels, 0 laughter) · 3 failed claims (excluded)



Transcript of test_00021

start-end (s) speaker [ground-truth label] text > outcome

0.3–8.1 S3 [turn] Okay, so let's get started. Uh, today's topic is the greenhouse effect, but before we dive in, did anyone have questions about the reading? > cut off

6.9–14.6 S2 [interruption] Oh—yeah! So, in chapter three they said carbon dioxide is like, this big blanket? But I wasn't sure how that actually, like, traps heat. > takes floor from S3

14.7–18.8 S2 [turn] Is it just bouncing the heat back to earth, or is it more... complicated than that? > yields to S5

18.8–23.6 S5 [turn] Honestly, that's such a good question. I always picture the blanket but it can't be that simple, right? > yields to S1

23.6–32.5 S1 [turn] Yeah, um, so, the analogy works up to a point. It absorbs infrared radiation and then re-emits it, so some of it goes back down? That's the general idea. > yields to S4

32.5–36.9 S4 [turn] Right, but the thing is, it's not blocking the sunlight, it's only acting on the heat leaving the earth.

35.6–38.5 S2 [failed claim] Oh, so like, the sunlight comes in but then the heat can't get out?

37.0–44.0 S4 [turn] —yeah! Exactly. The energy comes in as visible light, warms up the earth, and then the earth emits infrared. That's the stuff that gets snagged. > yields to S1

41.3–41.9 S5 [backchannel] mm-hm

41.6–41.9 S2 [backchannel] ohhh

44.3–49.9 S1 [turn] So the greenhouse gases kinda act like the glass in an actual greenhouse, just for a different kind of light. > yields to S2

49.8–53.2 S2 [turn] That makes way more sense. I kinda wish textbooks would just say that. > yields to S3

53.2–57.4 S3 [turn] Haha, right? They always try to get clever but it just gets confusing. > yields to S4

56.1–57.0 S5 [backchannel] Totally.

57.9–61.9 S4 [turn] Can we talk about methane though? They said it's way worse than CO2 but there's less of it. > yields to S1

62.1–65.7 S1 [turn] Yeah, methane's like—uh, I think over twenty times more potent? > yields to S3

64.2–65.3 S2 [backchannel] Wait, really? Wow.

65.7–71.1 S3 [turn] Yup, but it also breaks down faster than CO2, so that's why it's not the main focus. > yields to S5

71.0–74.2 S5 [turn] But isn't it still like, super bad for, um, short-term warming? > yields to S4

74.2–77.6 S4 [turn] Yeah, especially from stuff like cows and landfills. And like, fracking. > yields to S1

75.4–76.7 S2 [failed claim] Fracking, yeah. It's wild.

77.9–82.3 S1 [turn] There was that video with the flaming tap water, right? Methane leaking into wells. > holds, pauses

79.3–83.0 S4 [failed claim] Oh my gosh, yes, I remember that! That was from, like, that movie... what was it?

83.2–84.3 S1 [turn] —Gasland, I think. > yields to S5

84.3–86.0 S5 [turn] Yeah, Gasland! That was wild. > yields to S3

86.0–89.8 S3 [turn] Haha, I actually haven't seen that but now I feel like I have to. > yields to S2

89.7–91.3 S2 [turn] Oh you totally do, it's like, wow. > yields to S4

91.7–97.0 S4 [turn] Okay but, moving on, can we talk about the whole feedback loop thing? Like, permafrost and extra methane? > yields to S1

97.0–101.4 S1 [turn] Sure, so the permafrost stores a lot of methane, and if it melts, it releases even more greenhouse gases, which makes things worse. > cut off

100.4–102.6 S5 [interruption] That part actually stresses me out a little. > takes floor from S1

102.8–105.7 S5 [turn] But, um, I guess that's why there's so much urgency about it? > yields to S3

105.7–110.2 S3 [turn] Exactly. That's why all those scientists sound so freaked out sometimes. > yields to S2

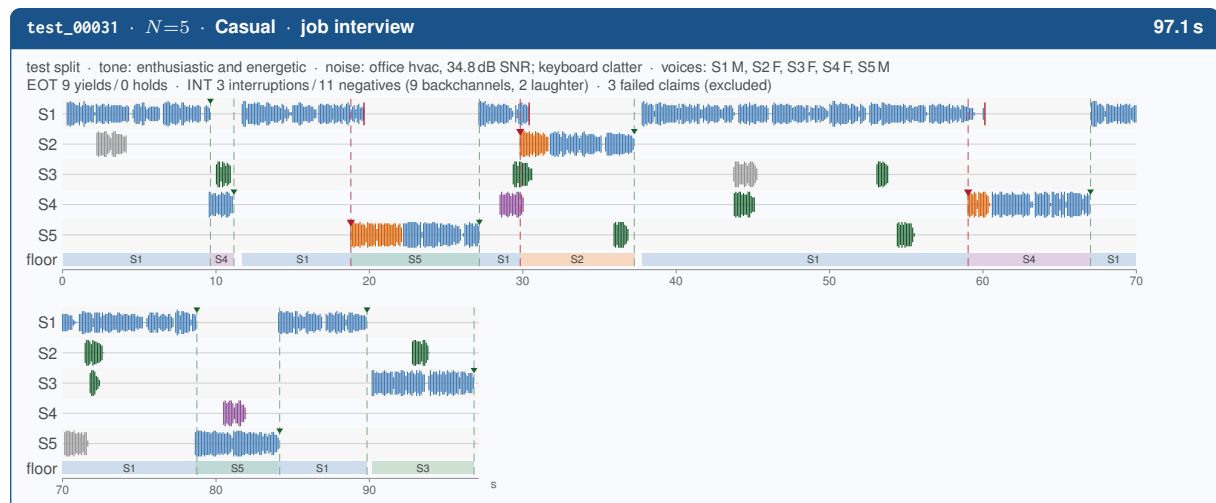
110.2–113.5 S2 [turn] Right, it's not just about warming, it's like, the domino effect. > yields to S4

113.5–115.5 S4 [turn] Yeah, once you start, it gets harder to stop. > yields to S1

116.0–121.3 S1 [turn] But, I mean, understanding it is the first step, so we're doing pretty well just discussing it like this. > yields, end

118.4–119.4 S5 [backchannel] Absolutely.

118.4–119.3 S2 [backchannel] Yeah, totally!



Transcript of test_00031

start-end (s) speaker [ground-truth label] text > outcome

0.3–6.4 S1 [turn] Hey everyone! Thanks for having me, I, uh, honestly couldn't believe I got the invite. Super excited to be here, really.

2.2–4.2 S2 [failed claim] Welcome, Ben, glad you made it—

6.6–9.6 S1 [turn] —uh, yeah, so should I just start introducing myself? Or...? > yields to S4

9.6–11.2 S4 [turn] Go ahead, we'd love that. > yields to S1

10.0–11.0 S3 [backchannel] Yeah, please do.

11.7–19.6 S1 [turn] Okay, cool. Uh, I'm Ben, I just wrapped up three years at Pinecreek—they're a little startup in logistics? I mostly wrangled data, lots of dashboards, lots of late nights with pizza. I really like figuring out how teams tick, you know? > cut off

18.8–22.2 S5 [interruption] Oh man, Pinecreek's that one with, like, the weird office dog, right? > takes floor from S1

22.2–27.2 S5 [turn] I think I saw them on TikTok with a basset hound in a conference room. That's you guys, right? > yields to S1

27.1–30.4 S1 [turn] Yeah! Chip. He snores through every all-hands. Like, full-on cartoon snoring. We all just gave up. > cut off

28.5–30.1 S4 [laughter] Haha that's adorable.

29.4–30.6 S3 [backchannel] That's hilarious.

29.8–31.7 **S2** [interruption] Alright, office dogs aside— **> takes floor from S1**

31.8–37.3 **S2** [turn] Can you talk about a project where, uh, you had to really improvise? Something that didn't go the way you pictured? **> yields to S1**

36.0–36.9 **S5** [backchannel] Yeah, curveballs.

37.8–49.9 **S1** [turn] Oh, for sure. Uh, last fall we, um, had this big client who changed their specs, like, a day before launch. Total panic. We had to rework half the dashboard overnight. I actually called in favors from people on support, just to test everything.

43.8–45.3 **S3** [failed claim] So what did you learn from that?

43.8–45.2 **S4** [backchannel] Wow, that's intense.

50.0–60.1 **S1** [turn] —it taught me to never assume requirements are final! But also, like, rallying everyone around one goal can really keep morale up, even at 2 a.m. I tried to make it kind of fun, actually. Sent memes. **> cut off**

53.1–53.9 **S3** [backchannel] Not wrong.

54.4–55.5 **S5** [backchannel] Memes fix everything.

59.0–60.5 **S4** [interruption] Can I just jump in— **> takes floor from S1**

60.6–67.0 **S4** [turn] —it sounds like you really value team energy. Can you tell us about, uh, your favorite thing about working with people? **> yields to S1**

67.1–77.3 **S1** [turn] Oh, absolutely. Honestly, it's just, like, riffing off each other's ideas? I love when someone throws out something wild and suddenly everyone's brainstorming. I feel like that's when magic happens.

70.1–71.7 **S5** [failed claim] Can I toss a question in real quick?

71.5–72.7 **S2** [backchannel] Yeah, same here.

71.8–72.4 **S3** [backchannel] Right.

77.4–78.8 **S1** [turn] —yeah, sure, go for it! **> yields to S5**

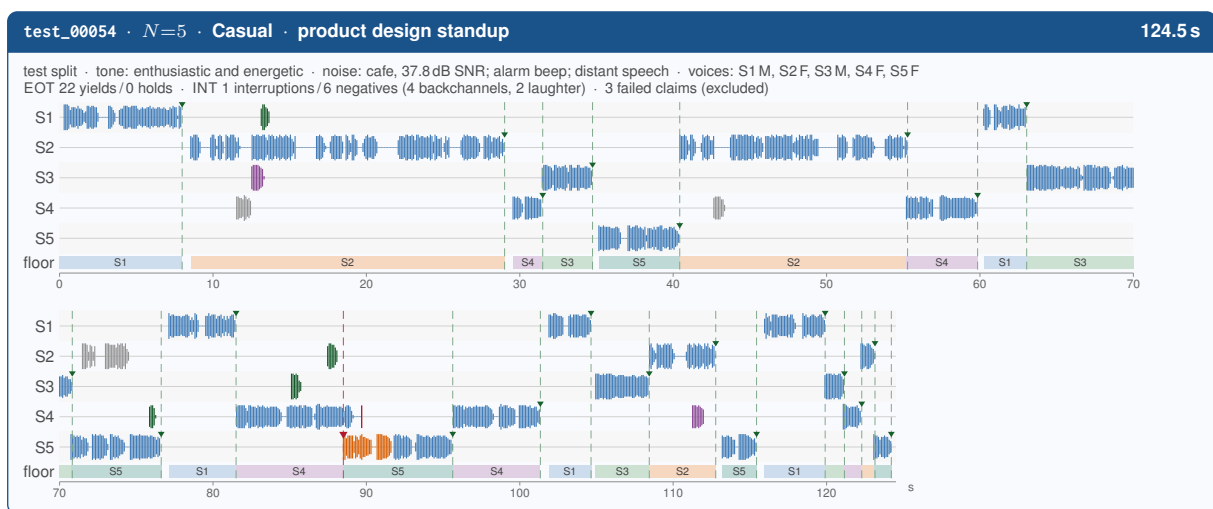
78.7–84.2 **S5** [turn] Uh, what's your favorite snack for late work nights? I feel like this is an essential cultural fit question. **> yields to S1**

80.5–82.0 **S4** [laughter] Haha, important.

84.1–89.8 **S1** [turn] Honestly? Microwave popcorn. Salted, never sweet. I can't do sugar at, like, 11 p.m. anymore. **> yields to S3**

90.2–96.8 **S3** [turn] Can't blame you. It's always popcorn or trail mix for me if I'm stuck late. I swear, popcorn's like, the universal programmer fuel. **> yields, end**

92.8–93.9 **S2** [backchannel] Totally agree.



Transcript of test_00054

start–end (s) speaker [ground-truth label] text **> outcome**

0.3–8.0 **S1** [turn] Alright, morning everyone, let's jump in. Amy, do you wanna kick us off with where we landed on the onboarding flow tweaks? **> yields to S2**

8.6–18.5 **S2** [turn] Yeah, sure. So, uh, not to brag but I did solve the weird loading spinner bug. I know, you're all welcome.

11.5–12.5 **S4** [failed claim] Ooooh, look at you!

12.5–13.3 **S3** [laughter] Haha, legend.

13.1–13.7 **S1** [backchannel] Right?

18.7–29.0 **S2** [turn] Yeah, yeah, applause later. So the screens now fade in instead of just, like, slapping the user in the face with white. **> yields to S4**

29.6–31.5 **S4** [turn] Thank god! The jump scare was real. **> yields to S3**

31.5–34.7 **S3** [turn] Seriously. That was waking me up faster than coffee on Monday. **> yields to S5**

35.2–40.4 **S5** [turn] I did notice it loads smoother, but the contrast on step two is still rough for accessibility. **> yields to S2**

40.4–45.9 **S2** [turn] Okay, yeah, that's valid. I didn't touch the color stuff yet.

42.7–43.4 **S4** [failed claim] Can I just—

46.0–55.3 **S2** [turn] —just wanted to get the animations feeling right first, but someone can totally pile on with palette suggestions. **> yields to S4**

55.2–59.8 **S4** [turn] I was gonna beg for brighter buttons. My eyes keep thinking they're disabled when they're not. **> yields to S1**

60.2–63.0 **S1** [turn] Noted. Jesse, you had some thoughts on copy, right? **> yields to S3**

63.1–70.8 **S3** [turn] Oh yeah, so the headline on the welcome page is still super stiff. I tried out 'Let's get you started!' instead of that spreadsheetsy nonsense. **> yields to S5**

70.7–73.2 **S5** [turn] I actually liked that. I think it's friendlier—

71.5–74.5 **S2** [failed claim] But not, like, too much confetti energy.

73.3–76.6 **S5** [turn] —right, exactly. Approachable but not, y'know, unhinged. **> yields to S1**

75.9–76.3 **S4** [backchannel] Mm-hm

77.1–81.5 **S1** [turn] Haha, yeah, no glitter bombs. Great. Who's got dashboard updates? Sarah? **> yields to S4**

81.5–89.7 **S4** [turn] So, um, dashboard got the new filter chips in. You can now filter by category and it actually remembers your settings when you come back. Magic! **> cut off**

85.1–85.8 **S3** [backchannel] Oh, sweet.

87.5–88.1 **S2** [backchannel] Finally!

88.5–91.7 **S5** [interruption] Does it persist across devices though? Or just the browser? **> takes floor from S4**

91.8–95.6 **S5** [turn] Because my test account forgot them when I swapped to mobile yesterday. **> yields to S4**

95.7–101.3 **S4** [turn] Oh! Uh, okay, that's just local storage right now. Cross-device is... a journey for next sprint. **> yields to S1**

101.9–104.6 **S1** [turn] Good catch, okay, we'll flag that for backlog. **> yields to S3**

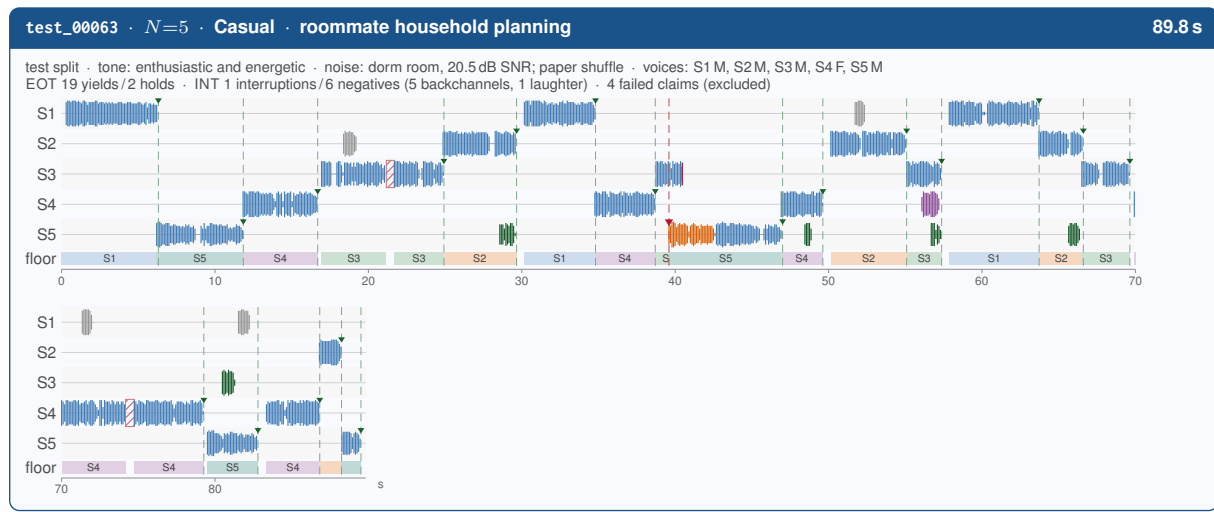
105.0–108.4 **S3** [turn] Hey, is the profile menu still glitchy for anyone else, or just me? **> yields to S2**

108.5–112.8 **S2** [turn] Yeah, mine popped open and then vanished like a ghost. **> yields to S5**

111.3–112.0 **S4** [laughter] Haha, same!

113.2–115.4 **S5** [turn] I saw it too. I'll log a bug for that. **> yields to S1**

116.0–119.9 **S1** [turn] Perfect, alright, anything else burning? Or can we free up our Friday? ▷ yields to S3
 119.9–121.2 **S3** [turn] No complaints here if we're done. ▷ yields to S4
 121.1–122.3 **S4** [turn] Let the weekend begin! ▷ yields to S2
 122.2–123.2 **S2** [turn] Later, team. ▷ yields to S5
 123.1–124.2 **S5** [turn] Bye all, see you Monday. ▷ yields, end

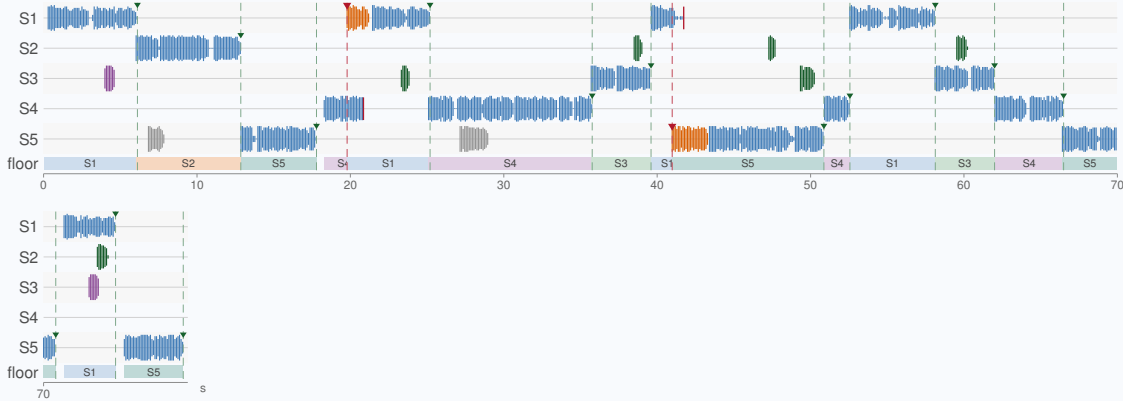


Transcript of test_00063

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.3 **S1** [turn] Alright y'all, so, uh, are we actually doing a chore wheel or are we pretending to be adults for like one more week? ▷ yields to S5
 6.2–11.8 **S5** [turn] Haha, dude, I vote chore wheel, let's get it over with. I can't keep stepping over laundry mountains in the hallway. ▷ yields to S4
 11.8–16.7 **S4** [turn] Oh my god, I know! The kitchen too. I keep finding like, mystery forks in my shoes. ▷ yields to S3
 16.9–21.2 **S3** [turn] Yeah, I'm in, but, uh, maybe we should pick days instead of jobs so it rotates? ▷ holds, pauses
 18.4–19.2 **S2** [failed claim] Wait, um—
 21.7–24.9 **S3** [turn] —so nobody's always stuck with like, trash or bathroom, you know? ▷ yields to S2
 24.9–29.7 **S2** [turn] Yeah, no, I was just gonna say, uh, what if someone can't do their day? Like, do we swap or just skip? ▷ yields to S1
 28.6–29.6 **S5** [backchannel] Ooh, good call.
 30.2–34.8 **S1** [turn] Yeah, like, if I have a late shift or something, I'm not scrubbing dishes at 2am, sorry. ▷ yields to S4
 34.8–38.7 **S4** [turn] Totally fair. So maybe we text the group to swap days if we need to? ▷ yields to S3
 38.8–40.5 **S3** [turn] Right, right, just like, post in the chat— ▷ cut off
 39.6–42.6 **S5** [interruption] I call next Sunday off! Haha, just putting that out there. ▷ takes floor from S3
 42.7–47.0 **S5** [turn] No, seriously, I have my cousin's wedding that weekend, so someone cover me? ▷ yields to S4
 46.9–49.6 **S4** [turn] I'll take your Sunday if you do my Tuesday, deal? ▷ yields to S2
 48.5–48.9 **S5** [backchannel] Deal!
 50.1–52.1 **S2** [turn] Haha, this is already complicated—
 51.7–52.4 **S1** [failed claim] Oh, come on—
 52.2–55.1 **S2** [turn] —but, like, maybe we set up a Google Sheet for it or something? ▷ yields to S3
 55.1–57.4 **S3** [turn] Yeah, spreadsheets are your love language, huh, Ben? ▷ yields to S1
 56.1–57.2 **S4** [laughter] Haha, nerd.
 56.7–57.3 **S5** [backchannel] Guilty.
 57.9–63.7 **S1** [turn] Okay so who's actually gonna make the sheet? Because if it's me, it's gonna look like a toddler's art project. ▷ yields to S2
 63.7–66.6 **S2** [turn] I'll set it up. It's, uh, kind of fun for me? Is that weird? ▷ yields to S3
 65.7–66.4 **S5** [backchannel] Not at all!
 66.5–69.6 **S3** [turn] Honestly, please do. You color-code things, it's beautiful. ▷ yields to S4
 69.9–74.2 **S4** [turn] Um, so, can we add plant watering to the list? Our fern is getting crispy. ▷ holds, pauses
 71.4–72.1 **S1** [failed claim] Oh, my bad—
 74.7–79.3 **S4** [turn] —also, whose turn is it to buy toilet paper? We've been out since, like, Tuesday. ▷ yields to S5
 79.5–82.8 **S5** [turn] Oh, that's on me. I'll do a Target run tonight so we don't suffer. ▷ yields to S4
 80.5–81.3 **S3** [backchannel] Lifesaver.
 81.5–82.3 **S1** [failed claim] You're a hero, man.
 83.4–86.8 **S4** [turn] Seriously, thank you. Okay, I feel way better about this now. ▷ yields to S2
 86.8–88.2 **S2** [turn] Same. Uh, team high five? ▷ yields to S5
 88.3–89.5 **S5** [turn] Woo! Adult solved. ▷ yields, end

test split · tone: thoughtful and contemplative · noise: living room tv distant, 36.7 dB SNR · voices: S1 M, S2 M, S3 F, S4 M, S5 F
EOT 14 yields / 0 holds · INT 2 interruptions / 8 negatives (6 backchannels, 2 laughter) · 2 failed claims (excluded)

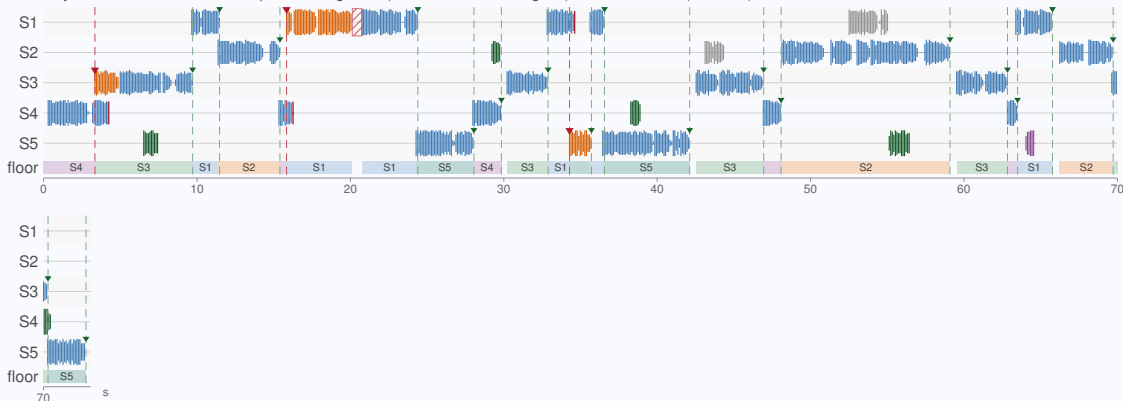


Transcript of test_00087

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.1 **S1** [turn] Alright, so, did anyone actually finish the book this time? Not pointing fingers, just... y'know, honest check-in. ▷ yields to S2
4.0–4.7 **S3** [laughter] haha
6.0–7.5 **S2** [turn] I did, uh, but it took me—
6.8–7.9 **S5** [failed claim] Oh, me too!
7.6–12.9 **S2** [turn] —a while. Like, honestly, I had to reread the first chapter twice. The writing... it's kind of dense, right? ▷ yields to S5
12.9–17.8 **S5** [turn] Yeah, totally. I had trouble getting through that middle section, the dream sequence part? ▷ yields to S4
18.3–20.9 **S4** [turn] Same here honestly, I, uh, got a bit lost who was— ▷ cut off
19.8–21.3 **S1** [interruption] Who was even talking, right? ▷ takes floor from S4
21.4–25.2 **S1** [turn] Like, he just drops you into someone's memories and expects you to follow along. ▷ yields to S4
23.3–23.9 **S3** [backchannel] mm-hm
25.1–30.7 **S4** [turn] Exactly! I kept flipping back. But there were moments I really liked, like the bit with the old photo album?
27.2–29.1 **S5** [failed claim] Oh, yeah, that was so bittersweet—
30.7–35.8 **S4** [turn] —and the way the author describes the grandmother's hands, did anyone else... like, visualize that? ▷ yields to S3
35.7–39.6 **S3** [turn] Yeah, I could really picture it. It reminded me a lot of my own grandma. ▷ yields to S1
38.5–39.0 **S2** [backchannel] Right.
39.6–41.7 **S1** [turn] Yeah, that whole scene hit hard. ▷ cut off
41.0–43.3 **S5** [interruption] I felt like—oh, sorry, can I jump in? ▷ takes floor from S1
43.4–50.9 **S5** [turn] I felt like the author was sort of using the photo album as a way to, I don't know, process grief? Did anyone else read it that way? ▷ yields to S4
47.3–47.8 **S2** [backchannel] mm-hm
49.3–50.3 **S3** [backchannel] Absolutely.
50.9–52.6 **S4** [turn] Yeah, that's how I felt too. ▷ yields to S1
52.6–58.1 **S1** [turn] Yeah, but do you think, like, it was too heavy-handed? I mean, sometimes it felt like the metaphors were a bit much. ▷ yields to S3
58.1–62.0 **S3** [turn] Well, I liked the language, but I get what you mean. There were spots where it tripped me up— ▷ yields to S4
59.5–60.2 **S2** [backchannel] Me too.
62.0–66.5 **S4** [turn] Yeah, mm, like the bit with the river and the clocks. I had to reread that, uh, three times. ▷ yields to S5
66.4–70.8 **S5** [turn] Haha, I actually thought that was beautiful, even if I didn't totally get it. ▷ yields to S1
71.3–74.7 **S1** [turn] Story of this book, huh—beautiful, confusing, all at once. ▷ yields to S5
73.0–73.7 **S3** [laughter] haha
73.5–74.2 **S2** [backchannel] Totally.
75.3–79.1 **S5** [turn] So, are we picking another one like this, or something lighter for next time? ▷ yields, end

test split · tone: professional and business-like · noise: office hvac, 30.4 dB SNR; alarm beep · voices: S1 F, S2 F, S3 F, S4 F, S5 F
EOT 19 yields / 1 holds · INT 3 interruptions / 6 negatives (5 backchannels, 1 laughter) · 2 failed claims (excluded)



Transcript of test_00140

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–4.2 **S4** [turn] Alright, team, are we locked in for the geography round? The timer's almost up— **> cut off**

3.3–4.9 **S3** [interruption] Wait, before you submit— **> takes floor from S4**

5.0–9.7 **S3** [turn] I think the capital of New Zealand is Wellington, not Auckland. Can we just double-check? **> yields to S1**

6.5–7.5 **S5** [backchannel] Oh, right.

9.7–11.5 **S1** [turn] Good catch! I always mix those up. **> yields to S2**

11.4–15.4 **S2** [turn] Auckland's the biggest city, but yeah, Wellington is the capital. She's right. **> yields to S4**

15.3–16.3 **S4** [turn] Got it, changing it— **> cut off**

15.8–20.1 **S1** [interruption] Wait, did anyone answer the question about the Nobel Prize winner from 2020? **> takes floor from S4 > holds, pauses**

20.8–24.4 **S1** [turn] I'm blanking on their name but it was the chemistry prize—CRISPR, right? **> yields to S5**

24.3–28.1 **S5** [turn] Yeah, Jennifer Doudna and Emmanuelle Charpentier. Super impressive. **> yields to S4**

28.0–29.8 **S4** [turn] Thanks, I never remember both names. **> yields to S3**

29.2–29.8 **S2** [backchannel] Same here.

30.2–32.9 **S3** [turn] Alright, moving on, the sports question is next. **> yields to S1**

32.8–34.6 **S1** [turn] Ooh, that's not my strong suit, so— **> cut off**

34.3–35.7 **S5** [interruption] I can take that one. **> takes floor from S1 > yields to S1**

35.6–36.6 **S1** [turn] —please do, haha. **> yields to S5**

36.5–42.1 **S5** [turn] Okay, it asks who won the women's World Cup in 2019. That's the US. I'll put that in. **> yields to S3**

38.3–39.0 **S4** [backchannel] Easy one!

42.5–44.0 **S3** [turn] Next is history—

43.1–44.4 **S2** [failed claim] Hang on, before we jump—

44.1–46.9 **S3** [turn] —to the next section, did we answer the pop culture one? **> yields to S4**

46.9–48.1 **S4** [turn] Oh, no, I don't think so. **> yields to S2**

48.1–53.9 **S2** [turn] Right, it was about the highest grossing film last year. It's probably 'Avatar: The Way of Water'.

52.5–55.1 **S1** [failed claim] Are we sure it beat 'Top Gun: Maverick'? I thought—

54.0–59.1 **S2** [turn] —yeah, I'm pretty sure 'Way of Water' pulled ahead in December. Box office numbers spiked. **> yields to S3**

55.1–56.5 **S5** [backchannel] That sounds right.

59.5–62.8 **S3** [turn] Alright, so, history question: Who signed the Magna Carta? **> yields to S4**

62.8–63.5 **S4** [turn] King John! **> yields to S1**

63.4–65.8 **S1** [turn] Haha, finally a medieval question I know. **> yields to S2**

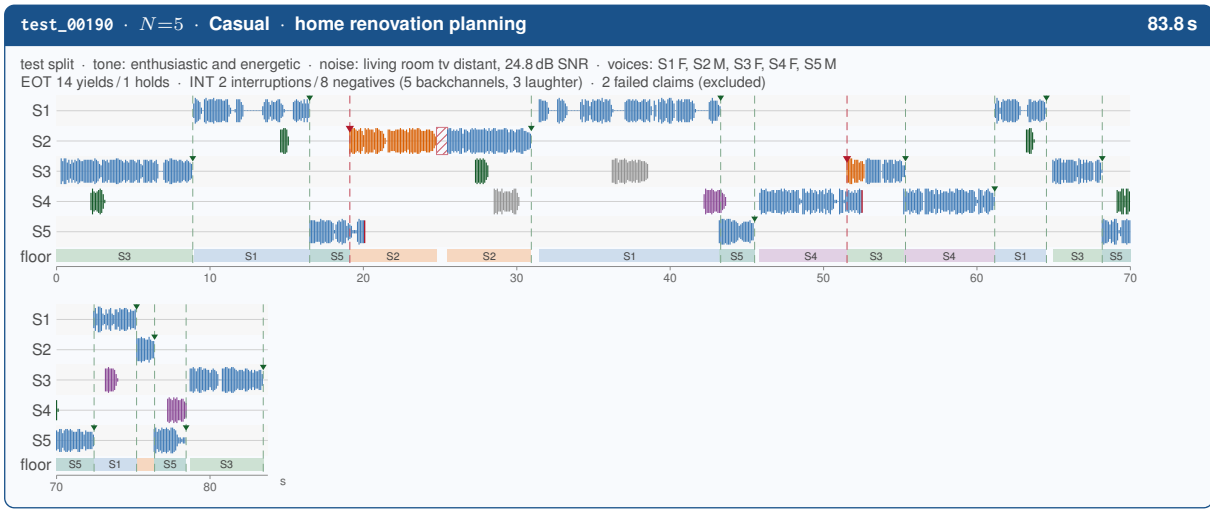
64.0–64.7 **S5** [laughter] haha

66.2–69.7 **S2** [turn] Okay, last one—science: What element has the symbol 'Fe'? **> yields to S3**

69.6–70.3 **S3** [turn] Iron. **> yields to S5**

70.0–70.5 **S4** [backchannel] Yep.

70.3–72.8 **S5** [turn] Alright, let's submit before the buzzer hits. **> yields, end**



Transcript of test_00190

start–end (s) speaker [ground-truth label] text **>** outcome

0.3–8.9 **S3** [turn] Okay, so kitchen first or living room first? I keep changing my mind, but the kitchen is just— it's a disaster. Like, the counter is literally foam. **> yields to S1**

2.2–3.2 **S4** [backchannel] Right, yeah.

8.9–16.5 **S1** [turn] Foam? I was going to say brittle cookie, but yeah. Soft, sticky, and not for food. **> yields to S5**

14.6–15.1 **S2** [backchannel] Oh, uh—

16.5–20.1 **S5** [turn] I mean, if the counter's that bad, you gotta start there. Nothing's more important than a solid counter. **> cut off**

19.1–24.8 **S2** [interruption] But, um, wait, what about the, uh, leak? Didn't you say there was a leak behind the walls in the living room? **> takes floor from S5 > holds, pauses**

25.5–30.9 **S2** [turn] Cause if you... like, if you do the kitchen and then you have to rip open the walls later, isn't that worse? **> yields to S1**

27.3–28.2 **S3** [backchannel] Yeah, good point!

28.5–30.2 **S4** [failed claim] Oh, that's so true. Mm-hm.

31.5–39.1 **S1** [turn] Honestly, either way, I vote we just move to Canada and leave the house for the raccoons.

36.2–38.6 **S3** [failed claim] Haha, okay, but actually, like, I kind of agree—

39.3–43.3 **S1** [turn] —they're better at home repairs than I am, let's be real. **> yields to S5**

42.2–43.7 **S4** [laughter] Haha, oh my god.

43.2–45.5 **S5** [turn] I'll bring my wrench to Canada if you need a hand. **> yields to S4**

45.8–52.5 **S4** [turn] So, um, do you want to prioritize what's critical? Like water issues always get worse. I lived through a ceiling collapse once. **> cut off**

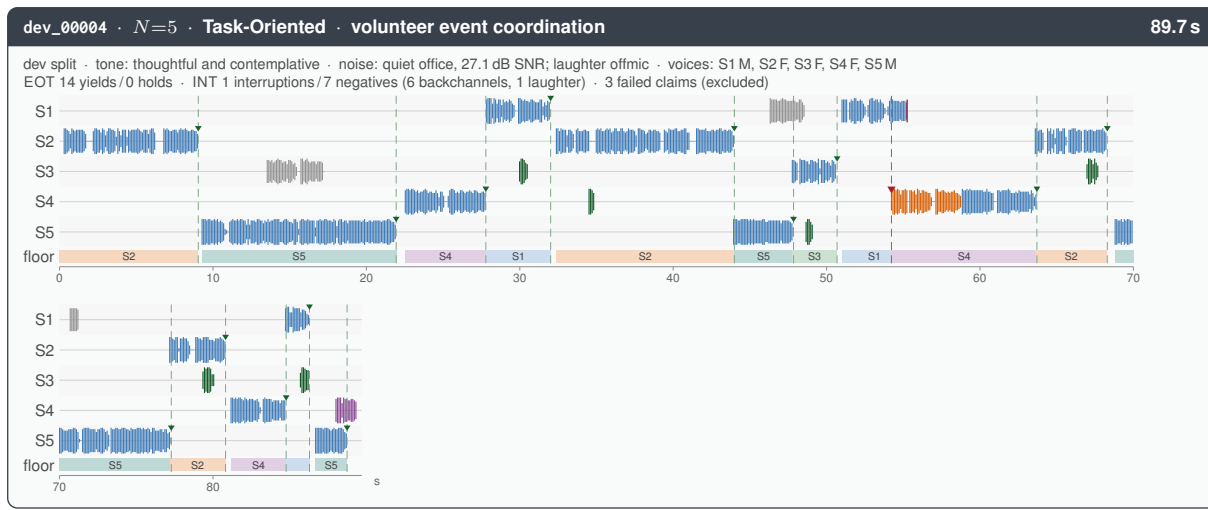
51.5–52.7 **S3** [interruption] Wait, what? Seriously? **> takes floor from S4**

52.8–55.3 **S3** [turn] You never told me that! Was it here or somewhere else? **> yields to S4**

55.2–61.2 **S4** [turn] This was back in college. My landlord just... ignored a drip until it became a waterfall. Wild times. **> yields to S1**

61.2–64.5 **S1** [turn] See, now the raccoon plan sounds less stressful. **> yields to S3**

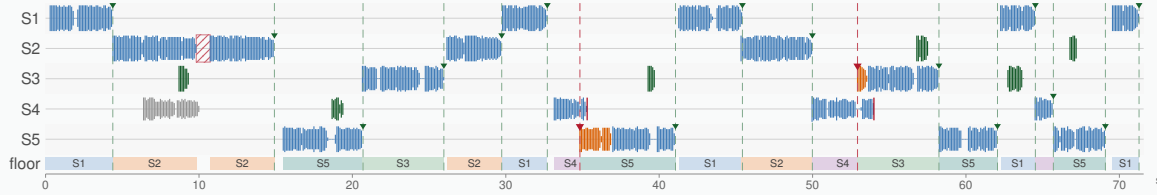
63.2–63.7 **S2** [backchannel] Uh-huh.
64.9–68.2 **S3** [turn] Okay, so... fix the living room leak first, then kitchen? Or— ▷ yields to S5
68.2–72.4 **S5** [turn] Yeah, drip first. Then you can actually cook without worrying your kitchen will float away. ▷ yields to S1
69.1–70.1 **S4** [backchannel] Exactly.
72.4–75.2 **S1** [turn] And if anyone's looking for a housewarming raccoon... ▷ yields to S2
73.2–74.0 **S3** [laughter] Haha, dibs!
75.2–76.4 **S2** [turn] I mean, they're handy, right... ▷ yields to S5
76.4–78.4 **S5** [turn] Just keep 'em outta the attic. ▷ yields to S3
77.2–78.5 **S4** [laughter] Haha, noted.
78.7–83.5 **S3** [turn] Alright, living room leak first. And then I want all of you painting cabinets with me, deal? ▷ yields, end



Transcript of dev_00004

0.3–9.0 **S2** [turn] Alright, let's focus in. Our biggest gap for Saturday's park cleanup is still the sign-in table, right? Has anyone spoken to Carla about supplies? ▷ yields to S5
9.3–15.6 **S5** [turn] Yeah, so I called her yesterday. She said she can drop off clipboards and the markers by eight, but she might not stick around to help set up.
13.5–17.2 **S3** [failed claim] Oh, but do we have enough name tags? Last event we almost ran out—
15.7–21.9 **S5** [turn] —she mentioned that too! She's bringing the leftover stack from last time, but if we think it'll be tight, I can swing by the store Friday night. ▷ yields to S4
22.5–27.8 **S4** [turn] How many volunteers do we actually expect, though? The RSVP list is kind of... well, it's vague. ▷ yields to S1
27.8–32.0 **S1** [turn] We're at, what, thirty-four confirmed? Plus the usual handful of walk-ins. ▷ yields to S2
30.0–30.5 **S3** [backchannel] Right
32.4–37.6 **S2** [turn] Let's budget for forty, just in case. We can always cut blank tags in half if we run low.
34.5–34.9 **S4** [backchannel] Hmm
37.7–44.0 **S2** [turn] Okay, switching gears—who's covering water and snacks? I don't want another situation like last May, you know? ▷ yields to S5
43.9–47.8 **S5** [turn] I'm happy to grab granola bars, but, um, does anyone have a big cooler for water? ▷ yields to S3
46.3–48.6 **S1** [failed claim] I do, but someone has to help me haul it—
47.8–50.7 **S3** [turn] Oh! I can meet you early and help carry it from your car. ▷ yields to S1
48.7–49.1 **S5** [backchannel] thanks
51.0–55.2 **S1** [turn] Perfect. So water's settled. What about first aid? I haven't seen the kit since last fall's event and we— ▷ cut off
54.2–58.8 **S4** [interruption] Sorry, but before that—do we have gloves for everyone? Last time people kept asking. ▷ takes floor from S1
58.9–63.7 **S4** [turn] I just don't want anyone turning up without them and then leaving early because their hands get wrecked. ▷ yields to S2
63.6–68.3 **S2** [turn] Good point. I'll check with Parks and Rec—they have extras, usually they just need notice. ▷ yields to S5
67.0–67.7 **S3** [backchannel] Sounds good
68.8–73.2 **S5** [turn] Great, let's recap the assignments before we wrap up, just so nothing falls through the cracks.
70.7–71.3 **S1** [failed claim] Hold up—
73.4–77.3 **S5** [turn] —did we ever decide who's doing the end-of-day trash tally for the city forms? ▷ yields to S2
77.2–80.8 **S2** [turn] I'll take that, it's fine. I can stick around after and fill them out. ▷ yields to S4
79.3–80.1 **S3** [backchannel] Perfect
81.2–84.8 **S4** [turn] Okay, that covers the main stuff. Anything else before we call it? ▷ yields to S1
84.7–86.3 **S1** [turn] Nope, that's all from me. ▷ yields to S5
85.7–86.3 **S3** [backchannel] Same
86.7–88.7 **S5** [turn] Looks like we're good then! See you all Saturday. ▷ yields, end
88.0–89.4 **S4** [laughter] haha, bright and early

dev split · tone: casual and friendly · noise: outdoor park, 34.2 dB SNR; notification chime · voices: S1 M, S2 F, S3 M, S4 F, S5 F
EOT 15 yields / 1 holds · INT 2 interruptions / 6 negatives (6 backchannels, 0 laughter) · 1 failed claims (excluded)

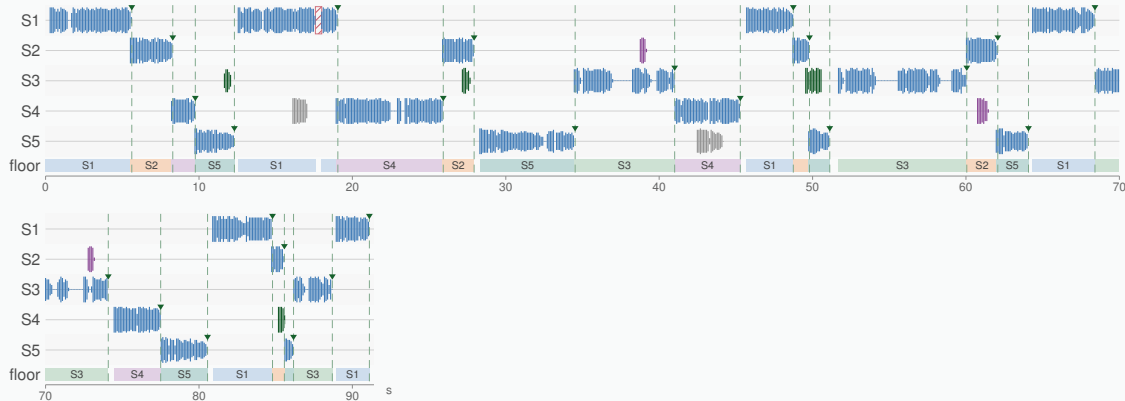


Transcript of dev_00029

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–4.4 **S1** [turn] Alright, everybody, thanks for coming. So, uh, let's dive in—living room first, right? ▷ yields to S2
 4.4–9.8 **S2** [turn] Right, if we're prioritizing budget, maybe we should look at floors before any big decorative stuff. ▷ holds, pauses
 6.4–10.0 **S4** [failed claim] But are we all actually sold on hardwood? I still think laminate might—
 8.7–9.3 **S3** [backchannel] Right
 10.7–14.9 **S2** [turn] —yeah, we can compare, just wanted the group's temperature on major costs first. ▷ yields to S5
 15.5–20.7 **S5** [turn] We can run numbers for both and see which fits the timeline. Hardwood install will take longer. ▷ yields to S3
 18.7–19.5 **S4** [backchannel] Exactly
 20.7–26.0 **S3** [turn] And cost-wise, laminate usually comes out a bit under half, depending on the finish. ▷ yields to S2
 26.2–29.7 **S2** [turn] But, uh, maintenance? I've had bad luck with chips in cheaper laminate. ▷ yields to S1
 29.8–32.7 **S1** [turn] Yeah, ours scratches if you just look at it wrong, honestly. ▷ yields to S4
 33.2–35.3 **S4** [turn] But the newer brands are tougher now, that's what the reviews say. ▷ cut off
 34.8–36.9 **S5** [interruption] Well, durability is one thing, but also— ▷ takes floor from S4
 37.0–41.1 **S5** [turn] —if there's any water exposure, hardwood will swell. That's a bigger repair. ▷ yields to S1
 39.3–39.8 **S3** [backchannel] Uh-huh
 41.3–45.4 **S1** [turn] So, sounds like we want to get samples of both. Shall I swing by the store tomorrow? ▷ yields to S2
 45.3–50.0 **S2** [turn] Yes please, and I can price out local installers to get some quotes ready by Friday. ▷ yields to S4
 50.0–54.0 **S4** [turn] I can help with the paint picking, after we land on the floor. Maybe blues for the living room? ▷ cut off
 52.9–53.6 **S3** [interruption] Wait— ▷ takes floor from S4
 53.7–58.2 **S3** [turn] Do we want to paint before or after the floors go in? Just thinking about mess. ▷ yields to S5
 56.8–57.6 **S2** [backchannel] Good question
 58.3–62.1 **S5** [turn] Definitely before, always. Less taping, less risk of damage. ▷ yields to S1
 62.3–64.5 **S1** [turn] Makes sense. So paint next week, floors after? ▷ yields to S4
 62.7–63.7 **S3** [backchannel] Sounds good
 64.5–65.7 **S4** [turn] Okay, I'm in. ▷ yields to S5
 65.7–69.1 **S5** [turn] Alright, I'll make a checklist and send it out tonight so we're all synced. ▷ yields to S1
 66.8–67.3 **S2** [backchannel] Awesome
 69.6–71.3 **S1** [turn] Teamwork, right? Thanks, everyone. ▷ yields, end

test split · tone: thoughtful and contemplative · noise: school hallway, 35.9 dB SNR; distant speech · voices: S1 F, S2 F, S3 F, S4 F, S5 M
EOT 25 yields / 1 holds · INT 0 interruptions / 7 negatives (4 backchannels, 3 laughter) · 2 failed claims (excluded)

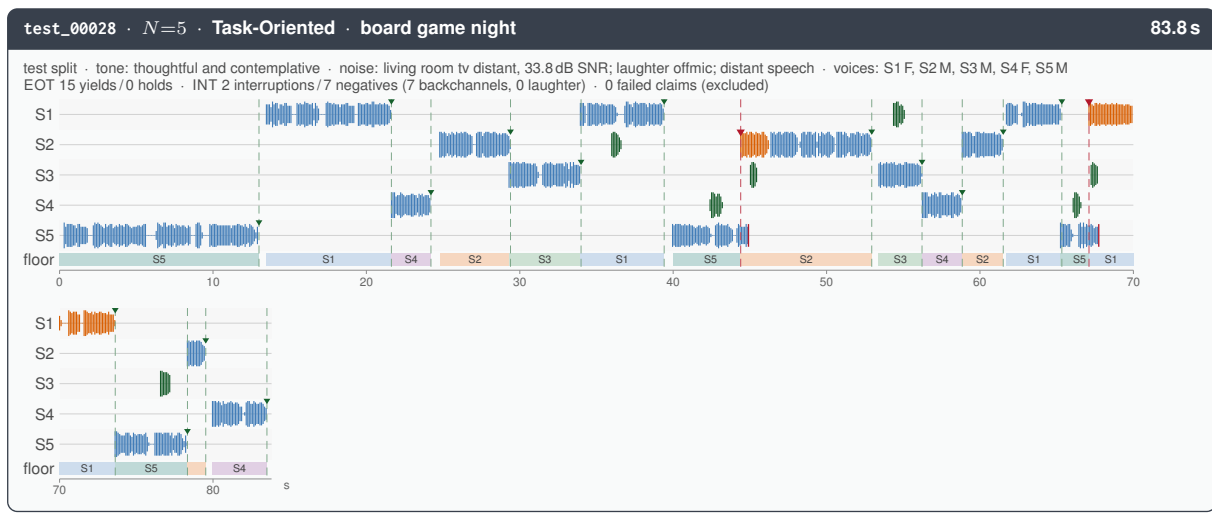


Transcript of test_00020

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–5.6 **S1** [turn] Alright, let's get started. First, just a quick check—we all saw the updated mockups I sent last night, right? ▷ yields to S2
 5.5–8.3 **S2** [turn] Yeah, they looked great by the way! I love the new icons. ▷ yields to S4
 8.2–9.8 **S4** [turn] Same here, they're super cute. ▷ yields to S5
 9.7–12.3 **S5** [turn] Absolutely, way more playful than the old ones. ▷ yields to S1
 11.7–12.1 **S3** [backchannel] mm-hm
 12.6–17.6 **S1** [turn] Thanks. So, the big decision: do we want to keep the sidebar navigation, or switch to the bottom nav bar? ▷ holds, pauses
 16.1–17.1 **S4** [failed claim] Oh, can I just—
 18.0–19.1 **S1** [turn] —yeah, go ahead. ▷ yields to S4
 18.9–25.9 **S4** [turn] Just, I feel the bottom nav would make it easier for mobile users? Like, my thumb's tired just thinking about the sidebar. ▷ yields to S2
 25.9–27.9 **S2** [turn] Totally agree, it's way more comfortable. ▷ yields to S5
 27.2–27.7 **S3** [backchannel] yeah
 28.3–34.5 **S5** [turn] Actually, I kinda like the sidebar because you can see more categories at once. Doesn't it help keep everything visible? ▷ yields to S3

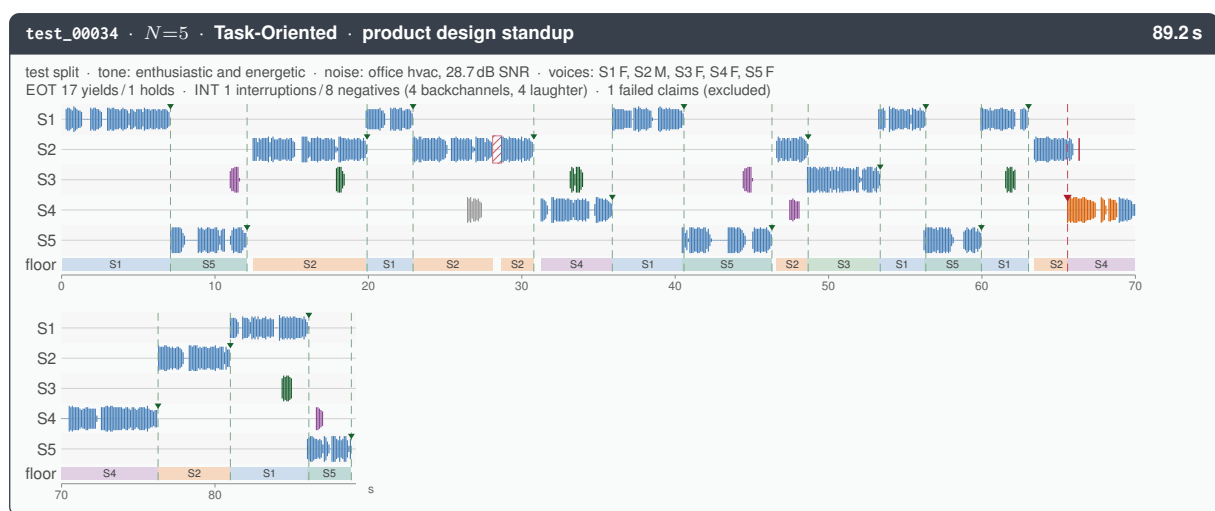
34.5–41.0 **S3** [turn] Maybe, but it also makes me scroll more. My index finger is not amused. $\triangleright$ yields to S4
 38.8–39.2 **S2** [laughter] haha
 41.0–43.2 **S4** [turn] Same! Sorry, Ben, but I'm team thumbs.
 42.5–44.2 **S5** [failed claim] Alright, alright, thumbs win—
 43.3–45.3 **S4** [turn] —but can we keep the sidebar for desktop? $\triangleright$ yields to S1
 45.7–48.7 **S1** [turn] Hybrid navigation is fine by me. Any strong objections? $\triangleright$ yields to S2
 48.7–49.8 **S2** [turn] Nope, works for me. $\triangleright$ yields to S5
 49.6–50.7 **S3** [backchannel] yeah, sounds good
 49.7–51.1 **S5** [turn] No arguments here. $\triangleright$ yields to S3
 51.7–60.1 **S3** [turn] Uh, can we talk about color palette though? I still feel like the new green is a little... radioactive. $\triangleright$ yields to S2
 60.0–62.1 **S2** [turn] Oh! Thank you, I thought it was just my monitor. $\triangleright$ yields to S5
 60.8–61.5 **S4** [laughter] haha same
 62.0–64.1 **S5** [turn] Wait, I like the green! It's kinda bold. $\triangleright$ yields to S1
 64.3–68.4 **S1** [turn] It's meant to pop, but if it's burning people's eyes, we can tone it down. Suggestions? $\triangleright$ yields to S3
 68.4–74.1 **S3** [turn] How about something a little more mossy and less, uh, nuclear waste? $\triangleright$ yields to S4
 72.8–73.2 **S2** [laughter] haha
 74.5–77.5 **S4** [turn] I'll pull a couple options to the Figma file after standup. $\triangleright$ yields to S5
 77.5–80.6 **S5** [turn] Perfect. Looking forward to seeing the less-glowing green. $\triangleright$ yields to S1
 80.9–84.8 **S1** [turn] Alright, sounds like we've got action items. Let's sync on the flows again tomorrow? $\triangleright$ yields to S2
 84.7–85.6 **S2** [turn] Works for me! $\triangleright$ yields to S5
 85.2–85.6 **S4** [backchannel] yeah
 85.6–86.2 **S5** [turn] I'm in. $\triangleright$ yields to S3
 86.2–88.7 **S3** [turn] I'll be here, radioactive or not. $\triangleright$ yields to S1
 88.9–91.1 **S1** [turn] Alright, thanks everyone. That's it for today. $\triangleright$ yields, end



Transcript of test_00028

start–end (s) speaker [ground-truth label] text $\triangleright$ outcome

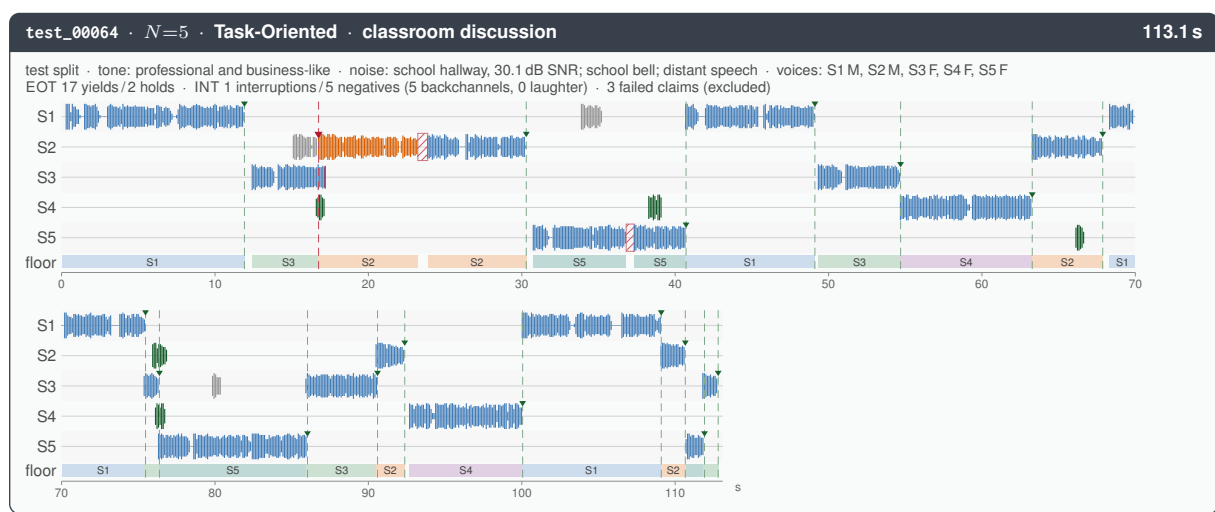
0.3–13.0 **S5** [turn] Okay, so before we deal, let's actually figure out which version of Settlers we're playing tonight. Are we good with just the base game or, um, does anyone want to add Cities and Knights or Seafarers? $\triangleright$ yields to S1
 13.5–21.6 **S1** [turn] Ooh, I vote for Seafarers! The boats are way more fun—that way we get to explore new islands and, you know, mess with everyone's plans a little. $\triangleright$ yields to S4
 21.6–24.2 **S4** [turn] I like Seafarers, too. It keeps things interesting. $\triangleright$ yields to S2
 24.8–29.4 **S2** [turn] Um, does Seafarers make it a lot longer though? I'm just, I don't know, worried it'll go late. $\triangleright$ yields to S3
 29.3–34.0 **S3** [turn] Not, uh, not if we skip the longest map. The four—islands scenario's pretty quick actually. $\triangleright$ yields to S1
 34.0–39.4 **S1** [turn] Yeah, and we can just set a timer if we're worried. Winner's whoever has the most points when it goes off! $\triangleright$ yields to S5
 36.0–36.6 **S2** [backchannel] Right
 40.0–44.9 **S5** [turn] Okay, so it sounds like Seafarers it is, shorter scenario, and we'll keep an eye on the time. Are there any objections before we start getting set up? $\triangleright$ cut off
 42.4–43.2 **S4** [backchannel] Sounds good
 44.4–46.2 **S2** [interruption] Um, just—wait, one more thing— $\triangleright$ takes floor from S5
 45.0–45.5 **S3** [backchannel] Nope
 46.4–52.9 **S2** [turn] I just, uh, I never remember: do you have to build the ship before you can move it, or is it placed and you can sail it right away? $\triangleright$ yields to S3
 53.4–56.2 **S3** [turn] Placed and sail on the next turn, if I'm remembering right. $\triangleright$ yields to S4
 54.4–55.1 **S1** [backchannel] Exactly
 56.2–58.8 **S4** [turn] You're right. New ships can't move until your next turn. $\triangleright$ yields to S2
 58.8–61.5 **S2** [turn] Okay, thanks. Sorry, rules always trip me up. $\triangleright$ yields to S1
 61.7–65.3 **S1** [turn] No worries! Half the fun is arguing about rules anyway, isn't it? $\triangleright$ yields to S5
 65.2–67.7 **S5** [turn] Haha, yeah, and I've got the rulebook here if we get really stuck. $\triangleright$ cut off
 66.1–66.6 **S4** [backchannel] Great
 67.1–73.6 **S1** [interruption] Wait, are we doing random board setup or the preset one? I, um, actually like the preset for Seafarers. $\triangleright$ takes floor from S5 $\triangleright$ yields to S5
 67.2–67.7 **S3** [backchannel] Mm-hm
 73.6–78.3 **S5** [turn] Let's do preset if everyone's okay. It's a little more balanced for the islands. $\triangleright$ yields to S2
 76.6–77.2 **S3** [backchannel] Agreed
 78.3–79.5 **S2** [turn] That works for me. $\triangleright$ yields to S4
 80.0–83.5 **S4** [turn] Alright, I'll start sorting out the tiles. Who wants to shuffle the cards? $\triangleright$ yields, end



Transcript of test_00034

start-end (s) speaker [ground-truth label] text > outcome

0.3–7.1 S1 [turn] Morning, everybody! Let's jump in—are we starting with the onboarding flow or did priorities shift again overnight? > yields to S5
7.1–12.1 S5 [turn] It's onboarding, unless the universe changed while I made coffee. > yields to S2
11.0–11.6 S3 [laughter] haha
12.5–19.9 S2 [turn] Uh, okay, so I did a quick pass on the welcome screen copy. I, um, tweaked the headline to be a little warmer, and... I think it reads more inviting now. > yields to S1
17.9–18.5 S3 [backchannel] Nice
19.9–22.9 S1 [turn] Can we see the new version? I wanna hype it up if it's good. > yields to S2
22.9–28.1 S2 [turn] Yeah, it's—uh, 'Welcome aboard, we're glad you're here.' It's, um, meant to be simple. I can share the preview— > holds, pauses
26.5–27.4 S4 [failed claim] Just to clarify—
28.7–30.8 S2 [turn] —I'll drop the mockup in Slack right now. > yields to S4
31.3–35.9 S4 [turn] Thanks. So do we want the greeting to sound personal, or just... friendly in general? > yields to S1
33.2–34.0 S3 [backchannel] Good point
35.9–40.6 S1 [turn] Oh, personal, please! Like, add a little sparkle. It's our first impression moment, ya know? > yields to S5
40.5–46.3 S5 [turn] Right, but no creepy sparkle. Too much sparkle and it's a clown car. > yields to S2
44.4–45.1 S3 [laughter] haha
46.6–48.7 S2 [turn] So a tasteful sprinkle of sparkle? > yields to S3
47.5–48.1 S4 [laughter] haha yes
48.6–53.4 S3 [turn] I like the changes, but—um, could we test it quickly with a new user? See their reaction? > yields to S1
53.3–56.3 S1 [turn] For sure, I can rope my roommate into a five-minute test. > yields to S5
56.2–60.0 S5 [turn] Will she demand compensation or just snacks? > yields to S1
59.9–63.0 S1 [turn] Oh, she accepts payment in leftover pizza. We're good. > yields to S2
61.5–62.2 S3 [backchannel] Perfect
63.4–66.3 S2 [turn] So, moving on... how are we handling the permissions screen? I know yesterday there was— > cut off
65.6–68.9 S4 [interruption] —Pushback from marketing about making it too, uh, intrusive. > takes floor from S2
68.9–76.3 S4 [turn] I had a chat with them. They're fine with us keeping the two-step prompt as long as the second screen explains why we need access. > yields to S2
76.3–81.0 S2 [turn] Okay, that makes things easier. Uh, I'll update the flow to show the rationale up front. > yields to S1
81.0–86.1 S1 [turn] Awesome! Also, microcopy for the rationale—I call dibs on writing something fun. > yields to S5
84.4–85.1 S3 [backchannel] Love it
86.0–88.9 S5 [turn] As long as it's not another pizza joke. > yields, end
86.6–87.1 S4 [laughter] haha



Transcript of test_00064

start-end (s) speaker [ground-truth label] text > outcome

0.3–11.9 **S1** [turn] Alright everyone, let's get started. Today we need to finalize the recommendations for the group project rubric, so maybe we can start by reviewing the draft that was circulated last night. ▷ yields to S3

12.4–17.2 **S3** [turn] I read through it this morning. Honestly, I think the section on collaboration needs tightening up. It's vague about what counts as effective teamwork. ▷ cut off

15.1–16.6 **S2** [failed claim] Yeah, totally agree with that—

16.6–17.2 **S4** [backchannel] mm-hm

16.8–23.2 **S2** [interruption] —could we maybe specify things like each member has to contribute at least one deliverable or... I dunno, something measurable? ▷ takes floor from S3 ▷ holds, pauses

23.9–30.3 **S2** [turn] I just think it's easier to grade if we're clear. Like, is attendance at team meetings enough, or are we expecting written input? ▷ yields to S5

30.8–36.8 **S5** [turn] That's a good question. Maybe we can add a line that participation means both being present and actively contributing— ▷ holds, pauses

33.9–35.3 **S1** [failed claim] Um, sorry to jump in—

37.3–40.7 **S5** [turn] —like, sharing documents or commenting on drafts, not just showing up. ▷ yields to S1

38.3–39.2 **S4** [backchannel] Exactly

40.7–49.1 **S1** [turn] Thanks, Maria. I wanted to ask if we want to actually require written evidence, or just keep it flexible for verbal contributions too? ▷ yields to S3

49.3–54.7 **S3** [turn] I'd vote for written evidence. Otherwise, it's way too easy for someone to claim they helped when there's no proof. ▷ yields to S4

54.7–63.3 **S4** [turn] That makes sense, but what about students who, um, aren't comfortable writing a lot? Should there be an option for audio messages or annotated slides maybe? ▷ yields to S2

63.3–67.9 **S2** [turn] Great idea! I mean, yeah, some people express themselves better talking than writing, right? ▷ yields to S1

66.1–66.7 **S5** [backchannel] Right

68.3–75.5 **S1** [turn] Okay, so we can say written, audio, or annotated slides count as evidence. Does that feel fair to everyone? ▷ yields to S3

75.4–76.4 **S3** [turn] Works for me. ▷ yields to S5

75.9–76.9 **S2** [backchannel] Absolutely!

76.1–76.8 **S4** [backchannel] mm-hm, yeah

76.3–82.4 **S5** [turn] It does. Also, should we talk about weighting? Like, how much collaboration's worth compared to the actual project content?

79.8–80.4 **S3** [failed claim] Wait—

82.4–86.0 **S5** [turn] —because last time, some people felt the peer assessment didn't count for enough. ▷ yields to S3

85.9–90.6 **S3** [turn] Sorry, just wanted to clarify—we're suggesting a specific percentage for collaboration, right? ▷ yields to S2

90.5–92.4 **S2** [turn] Yeah, like, is it twenty percent or...? ▷ yields to S4

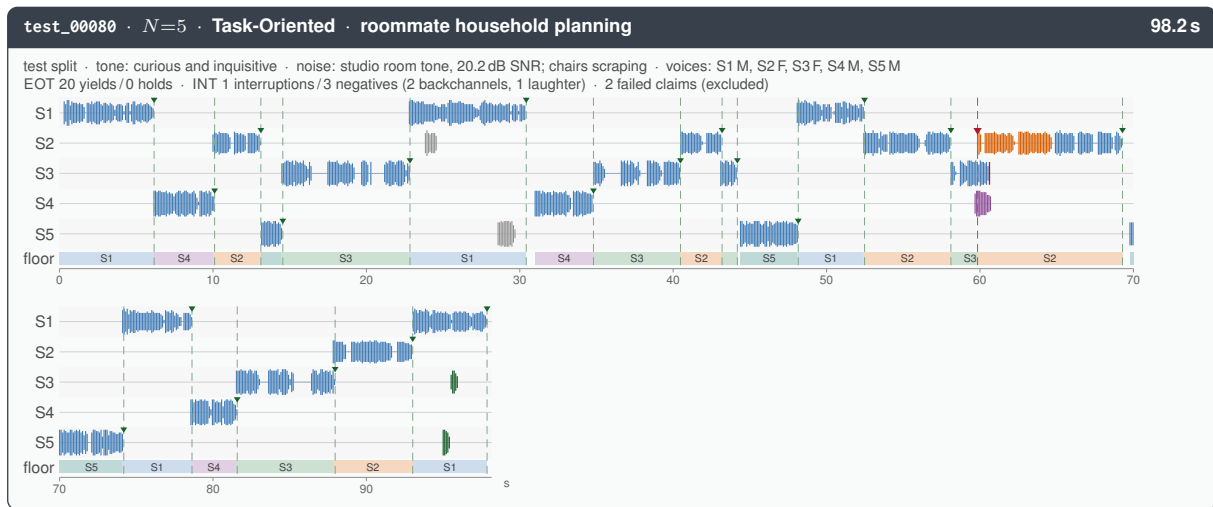
92.7–100.1 **S4** [turn] I think twenty is fair. That's enough to emphasize it's important but still leaves lots of room for the quality of the final product. ▷ yields to S1

100.0–109.1 **S1** [turn] Alright, let's put down twenty percent for collaboration, and the rest for content and presentation. Anything else before we move to deadlines? ▷ yields to S2

109.1–110.7 **S2** [turn] Nope, I think that covers it! ▷ yields to S5

110.7–111.9 **S5** [turn] Sounds good to me. ▷ yields to S3

111.8–112.8 **S3** [turn] All clear here. ▷ yields, end



Transcript of test_00080

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.2 **S1** [turn] Okay, crew meeting time! Uh, let's figure out the cleaning schedule for this month—like, what actually works for everyone? ▷ yields to S4

6.2–10.1 **S4** [turn] Um, so, are we talking, like, a chart with days, or just, uh, who does what? ▷ yields to S2

10.0–13.1 **S2** [turn] Chart! Definitely a chart. I want stickers. Or magnets. ▷ yields to S5

13.2–14.6 **S5** [turn] Oh, uh, magnets sound good... ▷ yields to S3

14.5–22.8 **S3** [turn] Can we get one of those whiteboards? Dry erase is less...permanent. I have commitment issues. ▷ yields to S1

22.8–25.2 **S1** [turn] Haha, okay, whiteboard with magnets. But, like—

23.9–24.6 **S2** [failed claim] —Wait, can I—

25.3–30.4 **S1** [turn] —who's actually available on weekends? Because last time it was just me and a mop for, like, forever. ▷ yields to S4

28.6–29.7 **S5** [failed claim] Yeah, uh, sorry about that.

31.0–34.8 **S4** [turn] I, um, actually have Saturday mornings open? Afternoon soccer is— ▷ yields to S3

34.8–40.5 **S3** [turn] Oh, same. But not before, like, ten a.m. Let's be reasonable. ▷ yields to S2

40.5–43.2 **S2** [turn] Haha, reasonable is like... noon for you, right? ▷ yields to S3

43.1–44.2 **S3** [turn] Rude. But fair. ▷ yields to S5

44.4–48.2 **S5** [turn] Um, I could do Sundays, if that helps, or maybe Fridays after work? ▷ yields to S1

48.1–52.5 **S1** [turn] Nice! Okay, so maybe Saturday: me, four, three. Sunday: five. Two, when are you free? ▷ yields to S2

52.4–58.1 **S2** [turn] Honestly, I'm kind of flexible, as long as it's not, like, before coffee. So, anytime after ten? ▷ yields to S3

58.1–60.6 **S3** [turn] Wow, guess we're all morning people now, suspicious. ▷ cut off

59.7–60.8 **S4** [laughter] Haha, right.

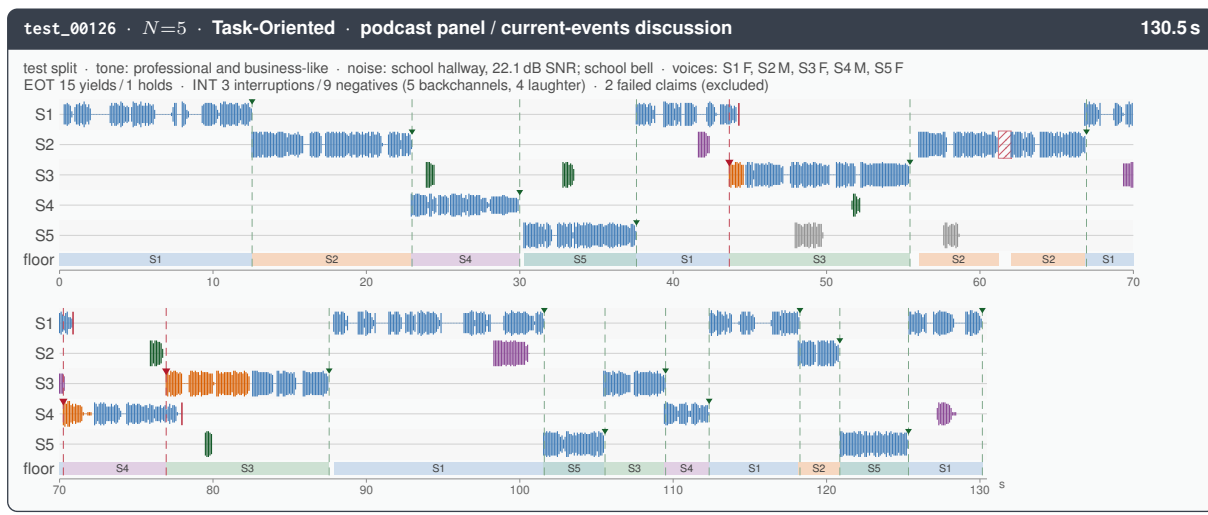
59.8–64.8 **S2** [interruption] Wait, do we need to assign specific chores or just show up and hope someone does the bathroom? ▷ takes floor from S3

64.9–69.3 **S2** [turn] Because every time it's, like, 'not it' for the toilet. I volunteer not? Kidding. ▷ yields to S5

69.8–74.2 **S5** [turn] Yeah, um, maybe, uh, we should rotate them? Like, a list that, uh, changes each week. ▷ yields to S1

74.1–78.6 **S1** [turn] Perfect! So we'll rotate, using, what, our magnet names? That way the whiteboard's... dynamic! ▷ yields to S4

78.6–81.6 **S4** [turn] And if someone can't make it, uh, just swap with someone else? ▷ yields to S3
 81.5–88.0 **S3** [turn] As long as it's not me, I'm allergic to toilets. Medically documented. ▷ yields to S2
 87.8–93.0 **S2** [turn] Haha, same. But, okay, I'll do whatever, as long as we get the board. I want to decorate it. ▷ yields to S1
 93.1–97.9 **S1** [turn] We have a plan! Magnets, whiteboard, rotation—no more mop solos. Let's make it happen. ▷ yields, end
 95.0–95.5 **S5** [backchannel] Yeah.
 95.5–96.0 **S3** [backchannel] mm-hm

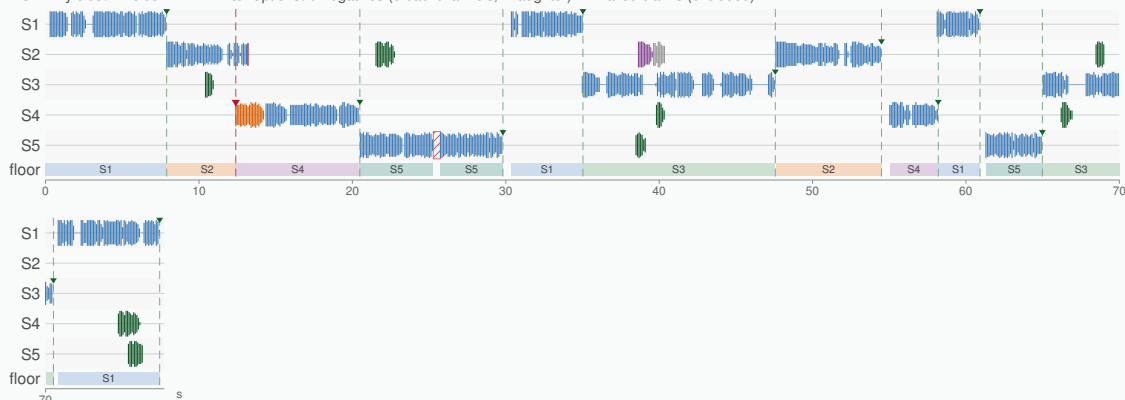


Transcript of test_00126

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–12.6 **S1** [turn] Alright, let's get started. Our first segment's housing price volatility, which, um, I'm sure no one has any strong opinions on at all. ▷ yields to S2
 12.5–23.0 **S2** [turn] I mean, if by strong opinions you mean utter confusion, then yes. But to frame it, national median home prices jumped nearly ten percent last quarter. So, what's really fueling that? ▷ yields to S4
 22.9–25.4 **S4** [turn] Interest rates are still a factor, obviously—
 23.9–24.5 **S3** [backchannel] Totally.
 25.5–30.0 **S4** [turn] —and on top of that, supply chain disruptions have made new builds slow to trickle in. ▷ yields to S5
 30.3–37.6 **S5** [turn] Yes, but I think it's not just supply. Demand is also up, especially from younger buyers who feel urgency now that remote work is here to stay. ▷ yields to S1
 32.8–33.6 **S3** [backchannel] Exactly!
 37.6–44.3 **S1** [turn] And let's not forget every hedge fund on the planet suddenly wants a slice. What could possibly go wrong there? ▷ cut off
 41.6–42.4 **S2** [laughter] Haha right.
 43.7–44.6 **S3** [interruption] Wait, can I just— ▷ takes floor from S1
 44.7–55.4 **S3** [turn] So I talked to two first-time buyers this week. Both lost four offers to all-cash bids. It's not just anecdotal—those institutional players are really swamping the entry level. ▷ yields to S2
 48.0–49.8 **S5** [failed claim] Yeah, we're seeing the same trend here.
 51.7–52.2 **S4** [backchannel] Mhm.
 56.0–61.2 **S2** [turn] But there's also a regional story, right? The Sunbelt cities are surging, but the Midwest is lagging. ▷ holds, pauses
 57.6–58.6 **S5** [failed claim] That's true, but—
 62.0–66.9 **S2** [turn] —that's not the whole picture, no, because remote work's also boosting some rural markets. ▷ yields to S1
 66.8–70.9 **S1** [turn] Rural broadband: finally the unlikely hero. ▷ cut off
 69.4–70.4 **S3** [laughter] Haha, exactly.
 70.2–72.1 **S4** [interruption] Let me jump in quickly— ▷ takes floor from S1
 72.3–78.0 **S4** [turn] There's a policy angle we shouldn't skip: local zoning restrictions are still a huge bottleneck for supply. Doesn't matter how much demand we have if cities keep blocking multifamily housing. ▷ cut off
 75.9–76.8 **S2** [backchannel] Absolutely.
 77.0–82.4 **S3** [interruption] Can I push back on that? Some cities tried reforms, but it takes ages for projects to get going— ▷ takes floor from S4
 79.5–80.0 **S5** [backchannel] Yep.
 82.6–87.6 **S3** [turn] —and by the time the units open, prices are even higher. It's like chasing a moving target. ▷ yields to S1
 87.9–101.6 **S1** [turn] So, to sum up, rampant demand, plus supply that's stuck in regulatory quicksand. Shall we pivot to solutions, or is doom-spiral analysis more fun? ▷ yields to S5
 98.3–100.6 **S2** [laughter] Haha, let's be bold and try for solutions.
 101.6–105.6 **S5** [turn] We've got about a minute left. So, each person: one change you'd make, rapid-fire. ▷ yields to S3
 105.5–109.5 **S3** [turn] Federal support for starter homes, make it easier for first-time buyers. ▷ yields to S4
 109.4–112.3 **S4** [turn] Zoning overhaul. Open it up, fast-track approvals. ▷ yields to S1
 112.4–118.3 **S1** [turn] Tax vacant lots so they get built. No more empty speculation. ▷ yields to S2
 118.2–120.9 **S2** [turn] Boost rental assistance for low-income families. ▷ yields to S5
 120.9–125.3 **S5** [turn] Nationwide push for prefab construction. It's fast, cost-effective, and scalable. ▷ yields to S1
 125.4–130.2 **S1** [turn] That's the closest we've come to optimism this whole hour. ▷ yields, end
 127.2–128.5 **S4** [laughter] Haha, fair.

test split · tone: thoughtful and contemplative · noise: cafe, 28.6 dB SNR; school bell, door knock · voices: S1 M, S2 M, S3 F, S4 F, S5 F
EOT 11 yields/1 holds · INT 1 interruptions/9 negatives (8 backchannels, 1 laughter) · 1 failed claims (excluded)

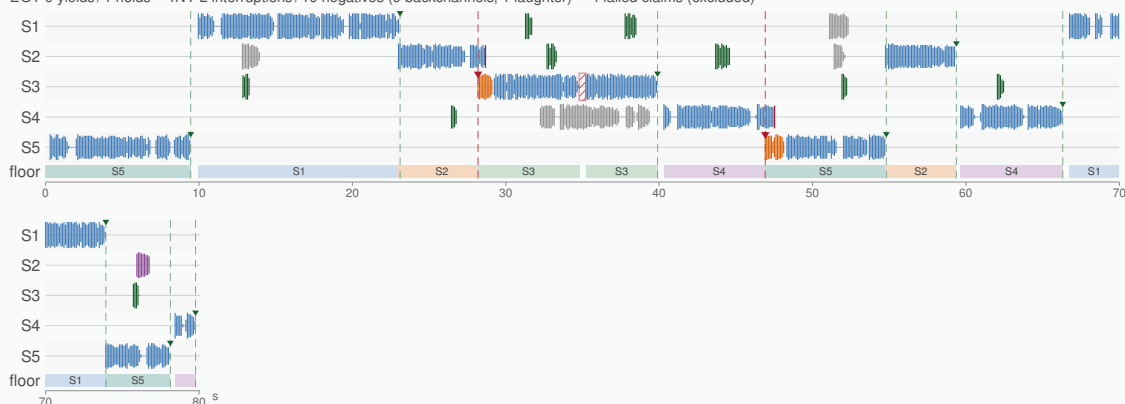


Transcript of test_00127

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–7.9 **S1** [turn] Good morning, everyone. Let's get started. Today I thought we could focus on the feedback for the prototype, especially the navigation flow. ▷ yields to S2
7.9–13.2 **S2** [turn] Yeah, honestly I was surprised by how much people struggled with that new menu. Like, uh, the nested tabs weren't as intuitive as we— ▷ cut off
10.4–10.9 **S3** [backchannel] Right.
12.4–14.2 **S4** [interruption] Excuse me, can I jump in? ▷ takes floor from S2
14.4–20.5 **S4** [turn] What if we flatten the structure? I mean, no more second-level tabs, just a single row. Would that solve most of it? ▷ yields to S5
20.5–25.3 **S5** [turn] It might. But did you see the feedback about discoverability? People missed whole features because, uh— ▷ holds, pauses
21.5–22.8 **S2** [backchannel] Yeah, exactly!
25.7–29.8 **S5** [turn] —they didn't scroll. So if we flatten, we risk making it even longer. That's my concern. ▷ yields to S1
30.4–35.0 **S1** [turn] Fair point. I wonder if a collapsible or accordion interface would help address both issues? ▷ yields to S3
35.0–41.2 **S3** [turn] Accordion menus scare users more than nested tabs, statistically speaking.
38.5–39.1 **S5** [backchannel] Yeah.
38.6–39.6 **S2** [laughter] Haha, yeah.
39.6–40.4 **S2** [failed claim] Okay but—
39.8–40.3 **S4** [backchannel] Mm-hm.
41.2–47.6 **S3** [turn] —but at least it would be, I don't know, less mysterious for new users. Maybe. ▷ yields to S2
47.6–54.5 **S2** [turn] What if we just add in-line hints or some little animations to draw attention, instead? Like, a subtle highlight when a new tab appears. ▷ yields to S4
55.0–58.2 **S4** [turn] We could A/B test that. We have the infrastructure for it. ▷ yields to S1
58.1–60.9 **S1** [turn] Agreed. Can someone volunteer to set that up this sprint? ▷ yields to S5
61.3–65.0 **S5** [turn] I can do it, provided someone else updates the onboarding flow copy. ▷ yields to S3
65.0–70.5 **S3** [turn] I'll take the copy updates. I like words more than accordion menus. ▷ yields to S1
66.2–67.0 **S4** [backchannel] Great.
68.5–69.1 **S2** [backchannel] Perfect.
70.8–77.4 **S1** [turn] Alright, let's recap: flatten structure gets a test, new hints in the UI, updates to onboarding. Looks like we're all set. ▷ yields, end
74.7–76.2 **S4** [backchannel] Yeah, sounds good.
75.4–76.4 **S5** [backchannel] Works for me.

test split · tone: thoughtful and contemplative · noise: living room tv distant, 33.0 dB SNR; notification chime, phone ring · voices: S1 M, S2 F, S3 M, S4 F, S5 M
EOT 9 yields/1 holds · INT 2 interruptions/10 negatives (9 backchannels, 1 laughter) · 4 failed claims (excluded)



Transcript of test_00164

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.5 **S5** [turn] Thank you all for joining. Let's begin by asking the candidate about challenges they've faced leading projects. Speaker one, would you like to start? ▷ yields to S1
10.0–17.9 **S1** [turn] Of course, thank you. So, could you describe a particularly challenging project you managed, and how you navigated conflicting team priorities?
12.8–14.0 **S2** [failed claim] That's a great framing—
12.9–13.3 **S3** [backchannel] uh-huh

18.0–23.1 **S1** [turn] —because in those situations, it’s really about prioritizing communication across roles. ▷ yields to S2

23.0–28.7 **S2** [turn] Absolutely! And could you walk us through, like, how you actually resolved disagreements? Maybe with a specific example? ▷ cut off

26.5–26.9 **S4** [backchannel] mm-hm

28.2–29.2 **S3** [interruption] If I can jump in— ▷ takes floor from S2

29.3–34.8 **S3** [turn] Sometimes it, uh, helps if you can tell us about a time when things didn’t work out and what you learned? ▷ holds, pauses

31.3–31.7 **S1** [backchannel] right

32.3–39.4 **S4** [failed claim] So, if I may, I’d also like to hear how you involve stakeholders during those disagreements. Do you, for instance—

32.7–33.4 **S2** [backchannel] Good point!

35.2–39.9 **S3** [turn] —yeah, um, maybe how you got buy-in from others who disagreed with your original plan? ▷ yields to S4

37.8–38.6 **S1** [backchannel] Exactly.

40.3–47.5 **S4** [turn] Thanks. Now, what I’m interested in is whether your approach leans more towards consensus-building, or if you tend to make executive decisions in a pinch? ▷ cut off

43.7–44.7 **S2** [backchannel] oh good question

46.9–48.2 **S5** [interruption] Let me just clarify— ▷ takes floor from S4

48.3–53.6 **S5** [turn] So, could you talk about a specific time you had to act quickly, maybe without consensus?

51.1–52.4 **S1** [failed claim] That’d be helpful, yes.

51.4–52.1 **S2** [failed claim] Sorry, just—

51.9–52.3 **S3** [backchannel] mm-hm

53.7–54.8 **S5** [turn] —Of course, go ahead. ▷ yields to S2

54.8–59.4 **S2** [turn] Just wanted to add, if you can tie it to handling pressure or time constraints, that’d be great. ▷ yields to S4

59.7–66.3 **S4** [turn] Thanks, yes. And maybe touch on how you debrief with the team afterward, especially if the decision wasn’t popular. ▷ yields to S1

62.0–62.5 **S3** [backchannel] right

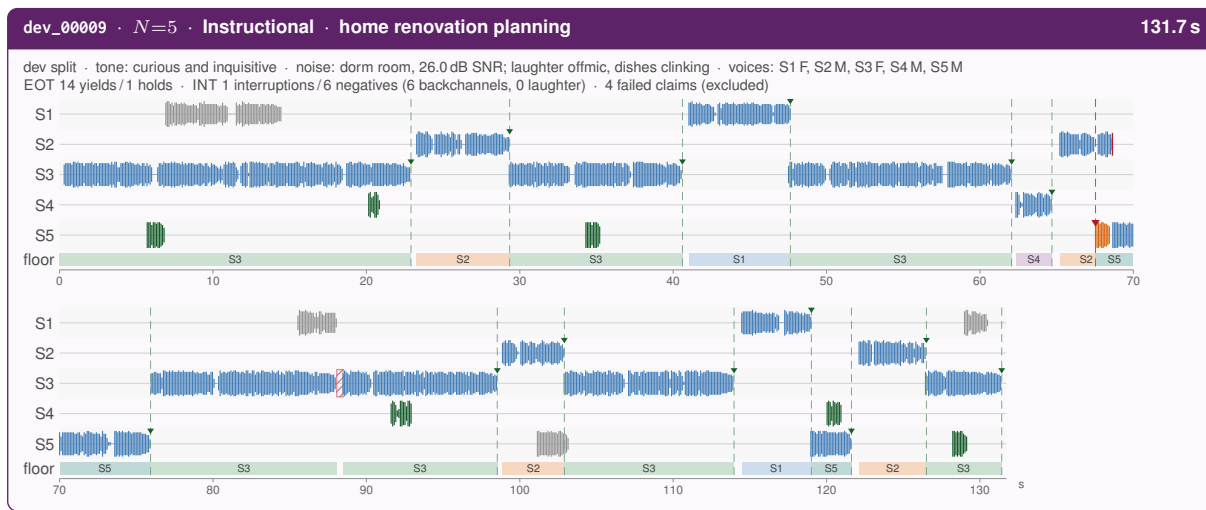
66.8–73.9 **S1** [turn] This is all very insightful. Thank you. I think we’ve covered most angles, unless anyone wants to ask a final, clarifying question? ▷ yields to S5

73.9–78.1 **S5** [turn] No, I believe that’s a wrap for this section. Great questions, everyone. ▷ yields to S4

75.7–76.1 **S3** [backchannel] yeah

76.0–76.8 **S2** [laughter] haha, totally

78.4–79.8 **S4** [turn] Agreed. Thank you all. ▷ yields, end



Transcript of dev_00009

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–10.6 **S3** [turn] All right, everyone, before we get too carried away, let’s walk through the renovation steps for the kitchen. We’ll start with the initial assessment and budgeting, then address priorities.

5.7–6.9 **S5** [backchannel] Yeah, sounds good.

6.9–14.5 **S1** [failed claim] But wait, before we really get into details, what exactly do you mean by “assessment”? Like, is that just looking around, or are we talking measurements—

10.7–22.9 **S3** [turn] —it’s a bit of both. You start by noting what needs repair, anything out of code, and yes, take full measurements as precisely as possible. You want a clear baseline before even thinking about style or colors. ▷ yields to S2

20.2–20.9 **S4** [backchannel] Makes sense.

23.3–29.3 **S2** [turn] So then, budgeting—are we talking a hard cap up front, or is it more like you adjust as you discover problems? ▷ yields to S3

29.3–40.6 **S3** [turn] Set a hard cap if you can, but always keep a buffer for surprises. Ten percent over your initial estimate is a good rule of thumb. And document everything—receipts, quotes, all of it. ▷ yields to S1

34.3–35.3 **S5** [backchannel] Oh, that’s smart.

41.0–47.6 **S1** [turn] But how do you even choose priorities? Like, we want new cabinets, but the wiring probably needs redoing too, and that’s way less fun. ▷ yields to S3

47.5–62.1 **S3** [turn] Safety and structure always come first. So electrical, plumbing, anything behind the walls? It has to be sorted before you even think about cabinets or finishes. Otherwise you’ll just end up tearing out your new stuff to fix something old later. ▷ yields to S4

62.3–64.7 **S4** [turn] Right, don’t want to redo drywall twice. ▷ yields to S2

65.2–68.6 **S2** [turn] Yeah, but what if you can only afford so much? Do you stage it out then, or just do the bare minimum? ▷ cut off

67.5–68.5 **S5** [interruption] Can I jump in— ▷ takes floor from S2

68.7–75.9 **S5** [turn] My friends ended up spreading their project over a couple of years, just tackling the essentials first. That way, you don’t burn out or go broke all at once. ▷ yields to S3

76.0–88.1 **S3** [turn] Exactly, phasing the work can make it way more manageable if budget’s tight. Just be sure your phases make sense structurally—don’t leave yourself stuck with half-finished wiring or things you can’t use in the meantime. ▷ holds, pauses

85.5–88.1 **S1** [failed claim] So, is there a checklist you follow? Or do you just—

88.5–98.5 **S3** [turn] —yes, always make a checklist. There are tons online, or you can make your own, but walk step by step—from demo, utilities, framing, up to finishes—so nothing gets missed. ▷ yields to S2

91.6–93.0 **S4** [backchannel] Yeah, checklists help.

98.9–102.9 **S2** [turn] I guess last thing—how do you pick contractors without just rolling the dice? ▷ yields to S3

101.2–103.2 **S5** [failed claim] References are everything, honestly.

102.9–114.0 **S3** [turn] You want to get at least three quotes, check references, actually call them. Ask to see previous work, and never pay more than a third up front. If it feels off, just move on—there are good ones out there. ▷ yields to S1

114.5–119.0 **S1** [turn] Cool, that's a lot more concrete than I expected. Thanks for laying it out so clearly. ▷ yields to S5

119.0–121.6 **S5** [turn] Yeah, this sounds way less intimidating now, honestly. ▷ yields to S2

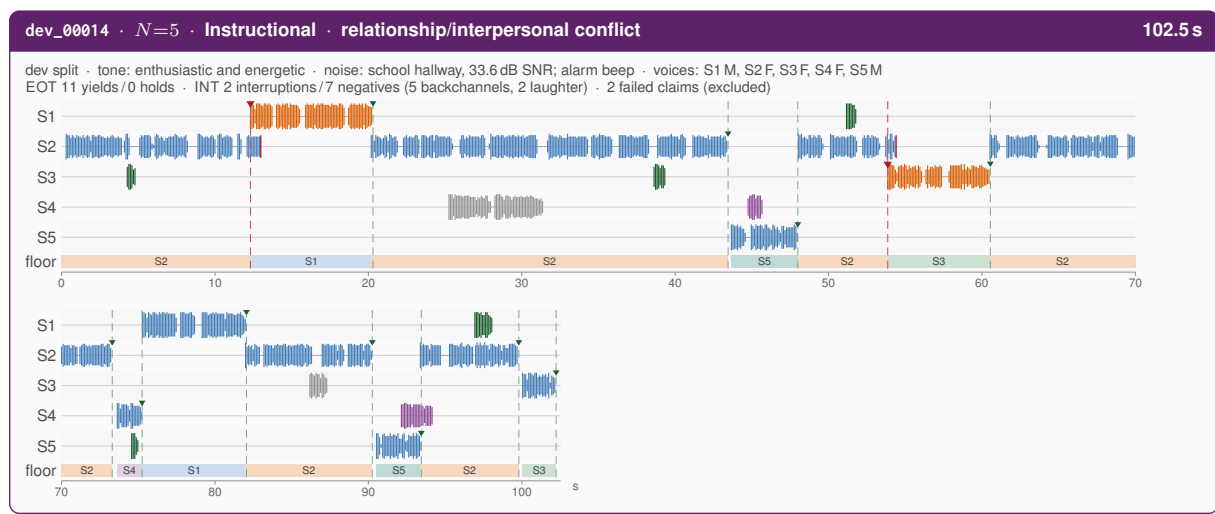
120.0–121.0 **S4** [backchannel] Yeah, agreed.

122.1–126.5 **S2** [turn] We should, uh, probably start by looking at those outlets behind the fridge, then. ▷ yields to S3

126.5–131.4 **S3** [turn] Perfect starting point. And let me know if you want those sample checklists—I'll send them over. ▷ yields, end

128.2–129.2 **S5** [backchannel] Awesome, thanks!

129.0–130.5 **S1** [failed claim] Great, really appreciate it.



Transcript of dev_00014

0.3–13.0 **S2** [turn] Alright, so today we're talking about conflict resolution in relationships, right? So, first off, the absolute key thing is active listening. I know it sounds simple, but actually doing it—like, really listening to the other person without planning your next comeback—makes a huge difference. ▷ cut off

4.3–4.9 **S3** [backchannel] Right!

12.3–20.3 **S1** [interruption] Wait, can I just clarify—when you say active listening, do you mean like paraphrasing what someone else said, or just not interrupting them? ▷ takes floor from S2 ▷ yields to S2

20.2–34.4 **S2** [turn] Good question! Both, actually. You want to not interrupt, but also reflect back what you understood. So, for example, you could say, 'I hear you're upset because I forgot your birthday, is that right?' That gives them a chance to correct you if you've missed something.

25.3–31.4 **S4** [failed claim] But what if they just get more annoyed when you repeat things back? I honestly feel like people can get super irritated if you—

34.6–43.5 **S2** [turn] —mm, yeah, I know what you mean. It can sound robotic if you're just parroting. The key is to be genuine with it, right? Use your own words, and keep your tone natural. ▷ yields to S5

38.6–39.4 **S3** [backchannel] Gotcha.

43.7–48.0 **S5** [turn] So, basically, don't sound like a guidance counselor from a TV show. ▷ yields to S2

44.8–45.7 **S4** [laughter] Haha, yeah.

48.0–54.4 **S2** [turn] Exactly! Sound like yourself. Another thing to remember—no blaming. Use 'I' statements. Like, 'I feel hurt when—' instead of, 'You always—' That changes the dynamic entirely. ▷ cut off

51.2–51.8 **S1** [backchannel] Uh-huh.

53.9–60.5 **S3** [interruption] Can I ask, what if the other person just, like, refuses to own up? You're using all the good techniques but they're just stonewalling? ▷ takes floor from S2 ▷ yields to S2

60.5–73.3 **S2** [turn] That's tough. Honestly, sometimes people need a break. If emotions are too high, suggest taking some time to cool off and revisit it later. And if someone really won't engage, you can only control your own side, you know? ▷ yields to S4

73.6–75.3 **S4** [turn] So, patience, basically. ▷ yields to S1

74.6–75.0 **S5** [backchannel] Mm-hm.

75.3–82.1 **S1** [turn] And is it better to bring things up right away, or give it a day? Because I feel like I'm always analyzing the timing. ▷ yields to S2

82.0–90.3 **S2** [turn] It really depends, but generally, you want to address things before resentment builds. Pick a time when you're both calm, but don't let it just sit forever. ▷ yields to S5

86.2–87.4 **S3** [failed claim] Oh, that makes sense.

90.5–93.5 **S5** [turn] So, don't schedule your big talk during rush hour. ▷ yields to S2

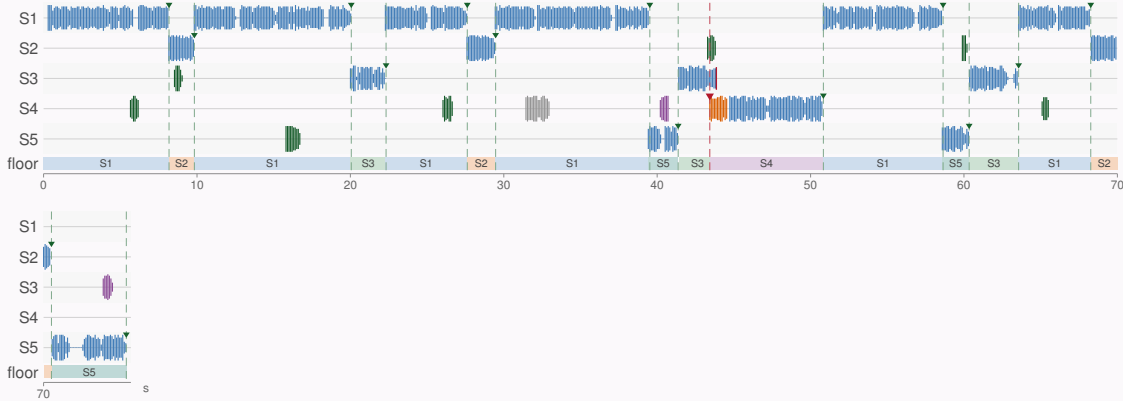
92.2–94.2 **S4** [laughter] Haha, yeah, or like right before bed.

93.4–99.8 **S2** [turn] Exactly, pick your moment. And honestly, practice helps, you get better every conflict you work through, not worse. ▷ yields to S3

96.9–98.1 **S1** [backchannel] Yeah, makes sense.

100.1–102.2 **S3** [turn] Thanks, this is actually super helpful. ▷ yields, end

test split · tone: curious and inquisitive · noise: studio room tone, 37.0 dB SNR; keyboard clatter · voices: S1 F, S2 M, S3 F, S4 F, S5 F
EOT 15 yields / 0 holds · INT 1 interruptions / 9 negatives (7 backchannels, 2 laughter) · 1 failed claims (excluded)



Transcript of test_00004

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.2 **S1** [turn] Alright, so the main thing I want to cover is how you handle an argument with someone you actually care about, instead of just, you know, avoiding it or letting it blow up. ▷ yields to S2

5.7–6.2 **S4** [backchannel] mm-hm

8.2–9.8 **S2** [turn] Like, actually talk about the thing, right? ▷ yields to S1

8.5–9.0 **S3** [backchannel] Yeah.

9.8–20.1 **S1** [turn] Right. So first, you gotta be clear what you're actually upset about. Don't come in halfway through, all passive-aggressive. State the thing—like, "I'm frustrated you didn't text me when you were late." Not, like, sulky silence. ▷ yields to S3

15.8–16.8 **S5** [backchannel] Revolutionary.

20.0–22.3 **S3** [turn] Is that before you cool off or after? ▷ yields to S1

22.3–27.6 **S1** [turn] After, always. If you're heated, you'll just say stuff you regret. Take a minute, breathe, figure out what the real thing is. ▷ yields to S2

26.0–26.7 **S4** [backchannel] oh yeah

27.6–29.5 **S2** [turn] So uh, what if they just shut down anyway? ▷ yields to S1

29.4–32.1 **S1** [turn] Good question. If they're clamming up, you can try saying—

31.4–33.0 **S4** [failed claim] Can you give like a script?

32.2–39.5 **S1** [turn] —yeah, sure. Try, "It feels like you're shutting down. Do you want to talk about it now, or can I check in later?" That gives them some space, but it doesn't drop it. ▷ yields to S5

39.4–41.4 **S5** [turn] So it's like, polite nagging. ▷ yields to S3

40.2–40.8 **S4** [laughter] haha

41.4–43.9 **S3** [turn] I mean, it sounds way less stressful than the cold war thing. ▷ cut off

43.3–43.8 **S2** [backchannel] For sure.

43.4–44.6 **S4** [interruption] Okay but— ▷ takes floor from S3

44.7–50.8 **S4** [turn] What if they're actually mad at you and won't say why? Sometimes you can tell, right, but if you ask, they just say "nothing". ▷ yields to S1

50.8–58.6 **S1** [turn] Oh, that's classic. Then you gotta be direct, but not confrontational. Like, "I can tell something's off, and I want to understand." Sometimes that's all it takes. ▷ yields to S5

58.6–60.3 **S5** [turn] Or you get the icy silence. ▷ yields to S3

59.9–60.3 **S2** [backchannel] yep

60.4–63.6 **S3** [turn] So basically just don't play games, and ask. Huh. ▷ yields to S1

63.6–68.3 **S1** [turn] Pretty much. And listen, even if it feels awkward. Most people just want to be heard, not fixed. ▷ yields to S2

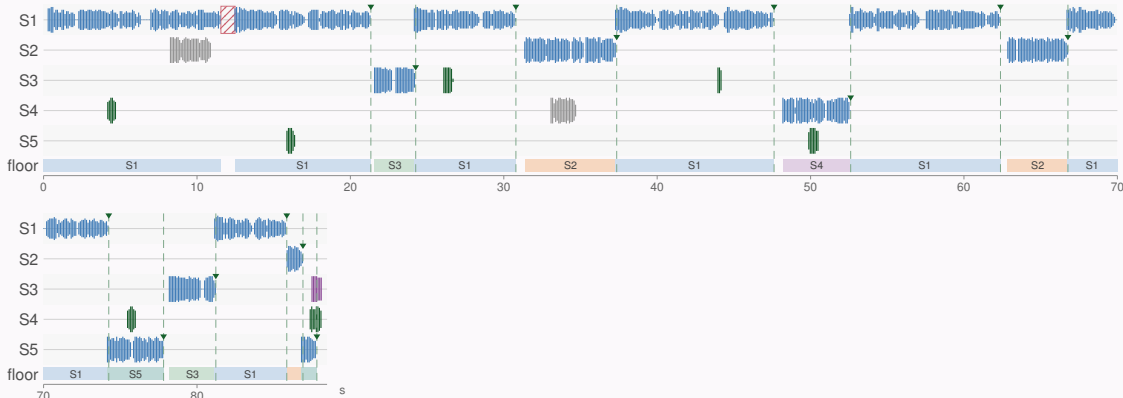
65.1–65.6 **S4** [backchannel] yeah

68.3–70.5 **S2** [turn] Man, wish I'd known that before last Thanksgiving. ▷ yields to S5

70.6–75.4 **S5** [turn] It's not too late. There's always next Thanksgiving disaster. ▷ yields, end

73.9–74.5 **S3** [laughter] haha

test split · tone: thoughtful and contemplative · noise: cafe, 35.2 dB SNR; phone ring, dishes clinking · voices: S1 M, S2 M, S3 F, S4 F, S5 F
EOT 14 yields / 1 holds · INT 0 interruptions / 8 negatives (7 backchannels, 1 laughter) · 2 failed claims (excluded)



Transcript of test_00019

start-end (s) speaker [ground-truth label] text > outcome

0.3–11.6 **S1** [turn] Alright! So this is called Wingspan. The goal is to build up a bird sanctuary with the prettiest, uh, collection basically. Each round you'll play birds, lay eggs, and use food to sort of power actions in your habitats. > holds, pauses

4.2–4.8 **S4** [backchannel] mm-hm

8.3–11.0 **S2** [failed claim] Oh—so, wait, do you only get one bird each turn?

12.5–21.3 **S1** [turn] —not quite, you actually pick one of four actions every turn. Playing a bird is just one option. You can also get food, lay eggs, or draw cards, depending where your cube goes. > yields to S3

15.9–16.4 **S5** [backchannel] right

21.6–24.3 **S3** [turn] I love the little eggs. Are they just for points, or...? > yields to S1

24.2–30.8 **S1** [turn] Haha, they're adorable, right? Eggs score points but they're also kinda like currency—you need them to play more birds in each row. > yields to S2

26.1–26.7 **S3** [backchannel] gotcha

31.4–35.2 **S2** [turn] So, uh, if you want to get more food, you said that's an action? How do you pick?

33.1–34.7 **S4** [failed claim] Like, is there a limit or—

35.4–37.4 **S2** [turn] —yeah, do you get stuck if you run out? > yields to S1

37.3–47.6 **S1** [turn] You each get one action cube per turn—that's how you pick. So, say you want food: you put your cube in the forest, then take food from the dice there. If the dice run low, they reroll the next time someone takes food. > yields to S4

44.0–44.2 **S3** [backchannel] ohhhh

48.2–52.6 **S4** [turn] That clears it up. And for the habitats—the three rows? Each is for a different action? > yields to S1

49.9–50.5 **S5** [backchannel] yeah

52.5–62.4 **S1** [turn] Exactly. Forest gives food, grasslands lay eggs, and the wetlands draw cards. And you can only place birds in a row if they match the habitat icons on their card—it's a little puzzle. > yields to S2

62.9–66.8 **S2** [turn] So, um, what happens if I can't pay the food or eggs for a bird I want? > yields to S1

66.7–74.2 **S1** [turn] Then you just wait, stock up on food or eggs with your other turns. There's no penalty, you just have to plan—it rewards being patient, honestly. > yields to S5

74.2–77.8 **S5** [turn] That feels comforting, actually. No rush, just... gentle strategy. > yields to S3

75.5–76.0 **S4** [backchannel] mm-hm

78.2–81.2 **S3** [turn] Alright, I'm ready to be chief egg collector! Can we start? > yields to S1

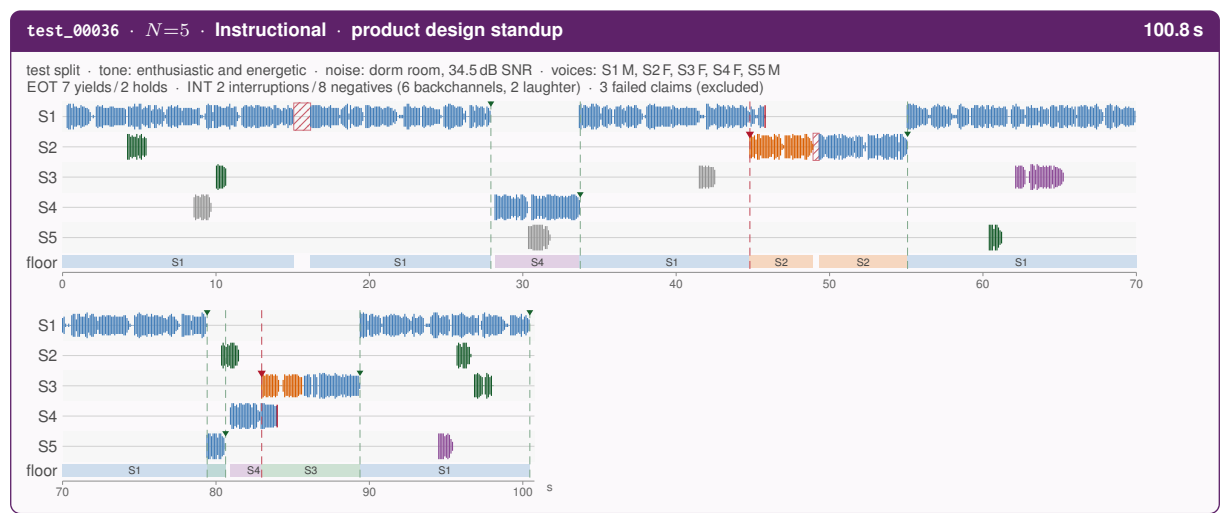
81.2–85.9 **S1** [turn] Haha, yes! Let's pass out the cubes and cards. I'll talk us through the first round as we go. > yields to S2

85.9–86.9 **S2** [turn] Sounds good. > yields to S5

86.8–87.8 **S5** [turn] I'm excited. > yields, end

87.4–88.2 **S4** [backchannel] me too

87.5–88.2 **S3** [laughter] haha right



Transcript of test_00036

start-end (s) speaker [ground-truth label] text > outcome

0.3–15.1 **S1** [turn] Alright team, let's dive in. I want to walk you through the new onboarding flow, so you can see exactly why we shifted the sequence and how that ties into our core interaction goals. First up, any time a new user signs up, they'll now hit the preferences survey right away instead of after profile setup. > holds, pauses

4.2–5.5 **S2** [backchannel] Oh, interesting.

8.6–9.7 **S4** [failed claim] Can you clarify—

10.0–10.7 **S3** [backchannel] That's new!

16.2–27.9 **S1** [turn] —because what we saw from user testing was, if they set preferences first, the system can tailor everything they see next. More relevant content, smoother experience. So, yeah, that's the big shift. Uh, any questions so far? > yields to S4

28.2–33.8 **S4** [turn] Yeah, how do we handle folks who skip the survey? Is there an escape hatch for the, uh, I-don't-want-to-give-you-data crowd? > yields to S1

30.4–31.8 **S5** [failed claim] I was wondering about that too.

33.7–45.8 **S1** [turn] Great catch. We added a 'skip for now' button, very visible, right below the survey. If someone skips, they get a generic dashboard and we nudge them again later. Actually, we're seeing about thirty percent take the survey up front, which honestly is much higher than the old flow. > cut off

41.5–42.6 **S3** [failed claim] Oh wow, that's awesome!

44.8–48.9 **S2** [interruption] Do you think that's because of the new placement, or is it just the nudge text? > takes floor from S1 > holds, pauses

49.3–55.1 **S2** [turn] Like, I'm curious if there's any evidence people feel more—uh, invested, I guess, with that up front? > yields to S1

55.1–64.3 **S1** [turn] Yeah, and that's the cool part. The numbers suggest it's both. It's immediate, but we also rewrote the text so it feels more like, 'help us help you,' not just a wall of questions.

60.4–61.3 **S5** [backchannel] That makes sense.

62.1–65.3 **S3** [laughter] Haha, yeah, walls of questions scare me off every time.

64.4–79.4 **S1** [turn] Totally. So, after preferences, they hit the guided profile setup, which is a step-by-step with progress markers. Speaker_5, you flagged last week that users drop out at the name-and-photo step. We created an option to add a photo later, and guess what? Drop-off dropped by fifteen percent. > yields to S5

79.4–80.6 **S5** [turn] Oh, that's really encouraging. > yields to S4

80.4–81.5 **S2** [backchannel] Nicely done.

80.9–84.0 **S4** [turn] So what's next after the profile setup? Is there an onboarding checklist or something? ▷ cut off

83.0–85.7 **S3** [interruption] Wait, is it gamified? Please tell me there's a badge. ▷ takes floor from S4

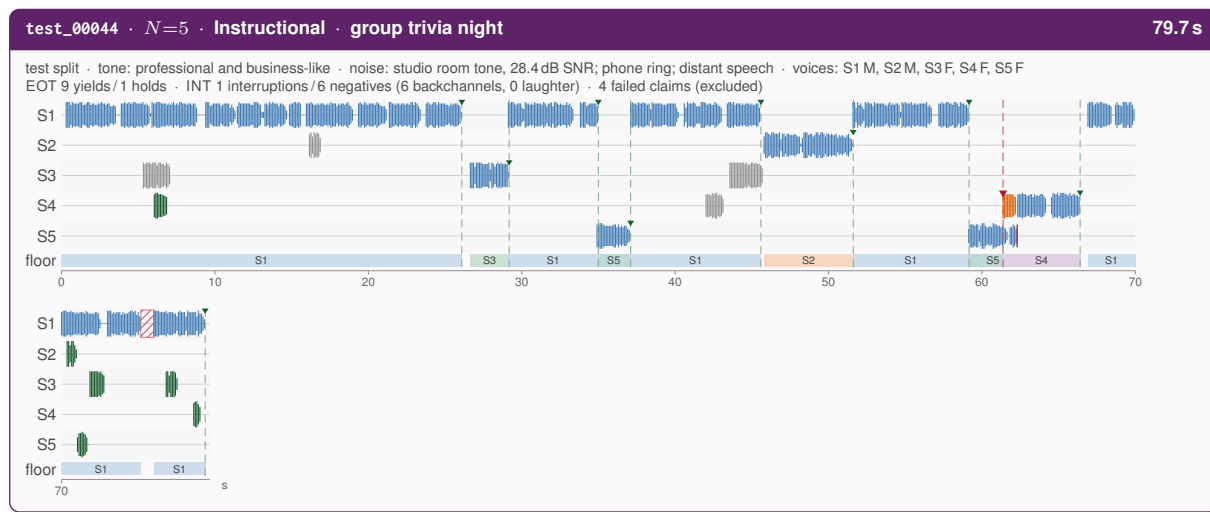
85.8–89.4 **S3** [turn] Because, I—I will absolutely do anything for a gold star, you know? ▷ yields to S1

89.4–100.5 **S1** [turn] Haha, not full badges yet, but we have a progress bar and confetti for first completion. We decided to keep it light, so no aggressive notifications, just a little celebration moment. More fun, less pressure. ▷ yields, end

94.5–95.5 **S5** [laughter] Haha, that's great.

95.7–96.6 **S2** [backchannel] Love that.

96.9–98.1 **S3** [backchannel] Confetti! Yes.



Transcript of test_00044

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–11.3 **S1** [turn] Alright, so for this round, we've got mostly geography questions. The format is that I'll read the question, then you'll have about a minute to discuss and write your answer on the sheet. I'll go through an example so we're all clear.

5.3–7.1 **S3** [failed claim] Oh perfect, I love geography stuff!

6.0–6.9 **S4** [backchannel] Yeah, same here.

11.5–21.2 **S1** [turn] So, for the example, if I ask, "What is the capital of Canada?" you can talk it out—maybe someone says Toronto, but someone else points out it's Ottawa. You can discuss before we finalize.

16.2–17.0 **S2** [failed claim] Wait, hang on—

21.4–26.1 **S1** [turn] —and then once you've agreed, just jot it down. You don't need to write an explanation, just the answer. ▷ yields to S3

26.6–29.2 **S3** [turn] Got it, so basically we talk first, then I write, right? ▷ yields to S1

29.1–35.0 **S1** [turn] Exactly, yeah. And only your first answer counts, so don't try to squeeze in a second one. Any questions about that? ▷ yields to S5

34.9–37.1 **S5** [turn] What happens if we spell something wrong? ▷ yields to S1

37.1–45.6 **S1** [turn] Minor spelling errors are fine as long as it's clear what you meant. For proper nouns, try to get as close as possible. They won't mark you down for a single missing letter. ▷ yields to S2

42.0–43.2 **S4** [failed claim] Phew, that's good to know!

43.6–45.7 **S3** [failed claim] Yeah, my handwriting gets messy when I'm rushing.

45.8–51.6 **S2** [turn] So, you said only the first answer counts—what if, like, we cross it out and write something else before you collect? ▷ yields to S1

51.6–59.2 **S1** [turn] Good question. As long as it's done before the host collects the sheets, your final written answer is what counts. No writing after the timer's up, though. ▷ yields to S5

59.1–62.3 **S5** [turn] Can we confer quietly while the timer's running, or does it have to be out loud? ▷ cut off

61.4–62.2 **S4** [interruption] Oh! Wait, so— ▷ takes floor from S5

62.3–66.4 **S4** [turn] Does whispering count as conferring? Or do we have to all discuss as a group? ▷ yields to S1

66.9–75.2 **S1** [turn] You can confer however you want—whisper, talk out loud, whatever's comfortable as long as everyone's included. The main thing is no phones or outside help. ▷ holds, pauses

70.4–71.0 **S2** [backchannel] Makes sense.

71.0–71.8 **S5** [backchannel] Clear.

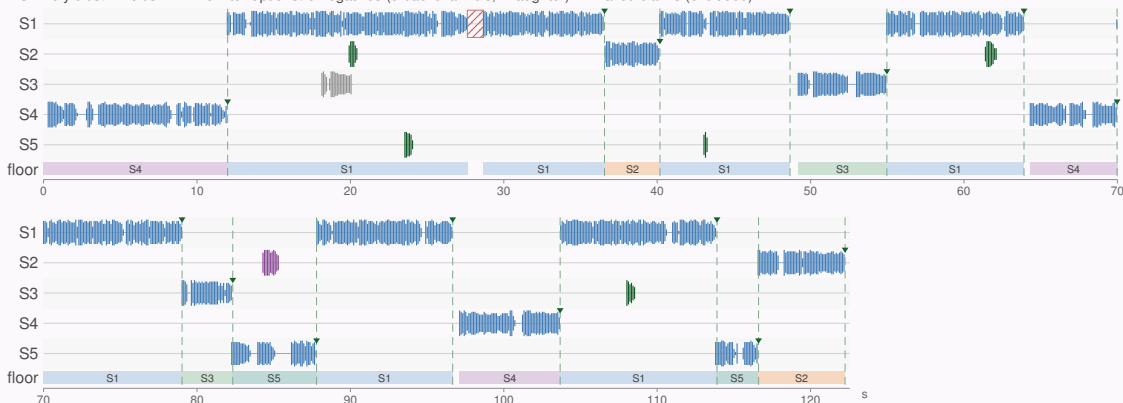
71.9–72.8 **S3** [backchannel] Right, no googling!

76.0–79.4 **S1** [turn] Alright, if there's nothing else, ready for the first official question? ▷ yields, end

76.8–77.6 **S3** [backchannel] Let's do it.

78.6–79.1 **S4** [backchannel] Yep!

test split · tone: curious and inquisitive · noise: outdoor park, 31.3 dB SNR; notification chime · voices: S1 M, S2 F, S3 F, S4 M, S5 F
EOT 15 yields / 1 holds · INT 0 interruptions / 6 negatives (5 backchannels, 1 laughter) · 1 failed claims (excluded)



Transcript of test_00114

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–12.0 **S4** [turn] Alright, everyone, let's dive in. Ben, you mentioned before the show you could walk us through how the offside rule works in soccer, which, frankly, I still get tripped up by sometimes. ▷ yields to S1

12.0–27.6 **S1** [turn] Uh, yeah, sure. So the basic idea is that a player is in an offside position if any part of their body, um, except the hands and arms, is nearer to the opponent's goal line than both the ball and the second-to-last defender, at the moment the ball's played to them. ▷ holds, pauses

18.1–20.2 **S3** [failed claim] Wait, so the ball and two defenders—

19.9–20.5 **S2** [backchannel] Mm-hm.

23.6–24.1 **S5** [backchannel] Right.

28.7–36.6 **S1** [turn] —yeah, so if you're closer to the goal than those two, you can be offside, but only if you're also involved in the play, like receiving the ball. ▷ yields to S2

36.6–40.2 **S2** [turn] So if a player is just standing there and not touching the ball, it doesn't count? ▷ yields to S1

40.2–48.7 **S1** [turn] Exactly. You have to be, uh, actively involved. The ref looks at whether you're interfering with play or gaining advantage simply by being there. ▷ yields to S3

43.0–43.3 **S5** [backchannel] Mm.

49.2–55.0 **S3** [turn] But what if, like, the defender steps forward at the last second? Is that how they do offside traps? ▷ yields to S1

55.0–63.9 **S1** [turn] Yeah, that's, uh, exactly it. Defenders time their step up so the attacker is suddenly ahead of them just as the pass is made, catching them offside. ▷ yields to S4

61.4–62.2 **S2** [backchannel] Got it.

64.3–70.0 **S4** [turn] Is it the assistant referee who decides, or the main ref? I always see them with the flag. ▷ yields to S1

70.0–79.0 **S1** [turn] Uh, the assistant referee raises the flag when they think there's an offside violation, but technically the center ref can overrule, though it's pretty rare. ▷ yields to S3

79.0–82.3 **S3** [turn] Right, I've seen a few controversial ones lately with VAR. ▷ yields to S5

82.3–87.8 **S5** [turn] VAR makes it... more controversial? Or just longer to watch? ▷ yields to S1

84.3–85.4 **S2** [laughter] Haha, yeah.

87.8–96.7 **S1** [turn] Both, honestly. They slow things down to review frame by frame, so even, uh, like a toe offside gets called. But theoretically it's more accurate. ▷ yields to S4

97.1–103.7 **S4** [turn] Okay, but what about if the defender deliberately plays the ball? I heard something about that changing offside? ▷ yields to S1

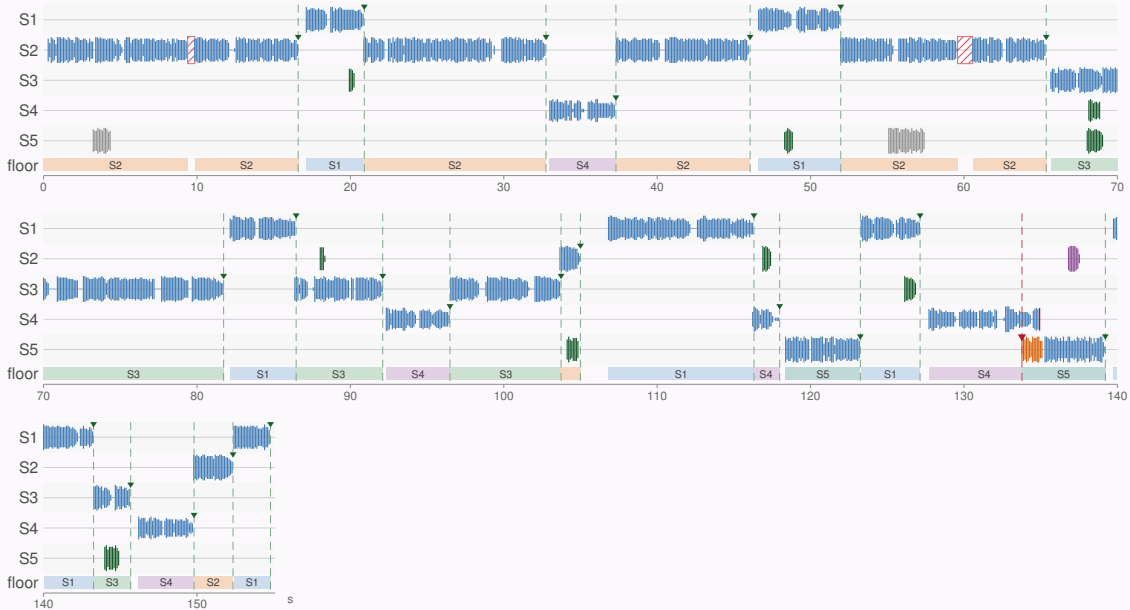
103.7–113.9 **S1** [turn] Yeah, so if a defender intentionally plays the ball, then the attacker can't be offside from that action, unless it's a save. It's, uh, newer guidance but that's the gist. ▷ yields to S5

108.0–108.6 **S3** [backchannel] I see.

113.8–116.6 **S5** [turn] I was today years old when I learned that. ▷ yields to S2

116.6–122.2 **S2** [turn] Thanks for breaking it down, Ben. I'm still not sure I could ref, but at least I won't yell at the TV for the wrong reason now. ▷ yields, end

test split · tone: enthusiastic and energetic · noise: studio room tone, 27.2dB SNR; dishes clinking · voices: S1 F, S2 F, S3 F, S4 M, S5 F
 EOT 23 yields / 2 holds · INT 1 interruptions / 10 negatives (9 backchannels, 1 laughter) · 2 failed claims (excluded)



Transcript of test_00128

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.4 **S2** [turn] Okay! So I know we said we'd swap recommendations, but I'm gonna gush about a book first, because I seriously just finished it last night and I can't stop talking about it. ▷ holds, pauses

3.2–4.4 **S5** [failed claim] Ooh, what is it?

9.9–16.6 **S2** [turn] 'Tomorrow, and Tomorrow, and Tomorrow.' Has anyone else read it yet? It's just—so good, like, ruined-me-for-other-books good. ▷ yields to S1

17.1–20.9 **S1** [turn] I've seen it in the bookstore. That's the one about the game designers, right? ▷ yields to S2

19.9–20.3 **S3** [backchannel] yeah

20.9–32.8 **S2** [turn] Yes! Sam and Sadie. It's this whole, like, decades-long friendship slash creative partnership, but it's not a romance—um, despite what you'd expect. And the way it talks about art and failure is so honest. ▷ yields to S4

33.0–37.3 **S4** [turn] Huh. Is it heavy on, uh, technical stuff? I mean, would someone who can't code get it? ▷ yields to S2

37.3–46.1 **S2** [turn] Oh, totally. You do not need to know a line of code. It's more about, like, ambition and vulnerability, and how you build something big with someone you care about. ▷ yields to S1

46.6–52.0 **S1** [turn] Is it one of those tear-jerker endings, though? Because I can't handle another 'A Little Life' situation. ▷ yields to S2

48.3–48.9 **S5** [backchannel] mm-hm

52.0–59.6 **S2** [turn] No no, I promise! It's... emotional for sure, but it's hopeful. Like, it made me cry in a cathartic way, not a destroy-your-week type way. ▷ holds, pauses

55.1–57.5 **S5** [failed claim] I love that. The characters sound awesome—

60.6–65.4 **S2** [turn] They're, like, so real, flaws and all. I finished and immediately wanted to text them, you know? ▷ yields to S3

65.7–69.0 **S3** [turn] So, are we just doing books or can I recommend a movie too?

68.0–69.1 **S5** [backchannel] Yeah, please.

68.1–68.9 **S4** [backchannel] Go for it.

69.1–81.7 **S3** [turn] Alright, 'After Yang.' I caught it last weekend and, honestly, it's one of the most quietly moving sci-fi films I've seen in ages. It's about a family and an android, but it's really about memory and grief. ▷ yields to S1

82.2–86.5 **S1** [turn] Is that the one with Colin Farrell? I remember a trailer but didn't see much buzz. ▷ yields to S3

86.4–92.1 **S3** [turn] Yep, that's it. It's super subdued, nothing explodes, but—I don't know, it just lingers with you. ▷ yields to S4

88.0–88.4 **S2** [backchannel] oohh

92.3–96.5 **S4** [turn] Does it, uh, have the whole dystopia vibe? Or is it more like, future-but-still-warm? ▷ yields to S3

96.5–103.7 **S3** [turn] Future-but-still-warm, exactly. There's technology, but the focus stays on the people. And it's only ninety minutes, so— ▷ yields to S2

103.7–105.0 **S2** [turn] Sold. Adding it to my list. ▷ yields to S1

104.1–104.9 **S5** [backchannel] me too

106.8–116.3 **S1** [turn] Alright, my turn. If you want something short and weird, read 'We Have Always Lived in the Castle' by Shirley Jackson. It's unsettling but, uh, not scary exactly. Just beautifully creepy. ▷ yields to S4

116.2–118.0 **S4** [turn] Classic. Good choice. ▷ yields to S5

116.9–117.4 **S2** [backchannel] oh right

118.4–123.2 **S5** [turn] I haven't read that one yet—does it have, like, horror-horror or is it more psychological? ▷ yields to S1

123.3–127.1 **S1** [turn] Psychological. You'll just feel uneasy, like you left a door unlocked. ▷ yields to S4

126.1–126.9 **S3** [backchannel] oh that's good

127.7–134.9 **S4** [turn] Alright—I'll throw in a wild card. If you want a cozy movie, try 'The Lunchbox.' It's a slow-burn, Indian film, all about loneliness and hope, with food tying everything together. ▷ cut off

133.8–135.2 **S5** [interruption] I loved that one! ▷ takes floor from S4

135.3–139.2 **S5** [turn] It made me want to write handwritten notes to strangers and start cooking again. ▷ yields to S1

136.8–137.6 **S2** [laughter] haha same

139.7–143.3 **S1** [turn] You all are making my to-watch list way too long. But this is good. ▷ yields to S3

143.3–145.7 **S3** [turn] So we're all set for like, the next month, right? ▷ yields to S4

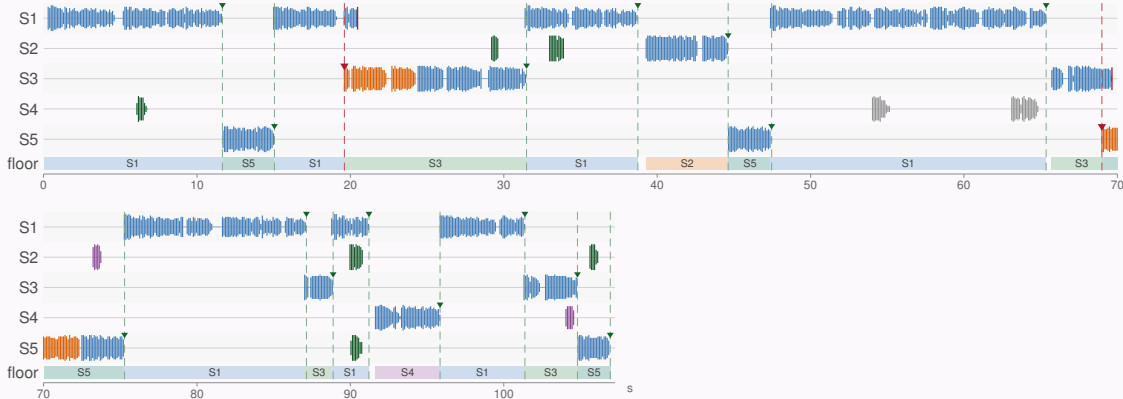
144.0–144.9 **S5** [backchannel] absolutely

146.2–149.8 **S4** [turn] At least. And hey, next time, maybe we do a theme night or something. ▷ yields to S2

149.8–152.3 **S2** [turn] I'm in! Someone start a group chat for all the reviews. ▷ yields to S1

152.4–154.8 **S1** [turn] Done. I'll send an invite after we wrap up. ▷ yields, end

test split · tone: curious and inquisitive · noise: office hum, 22.0 dB SNR; keyboard clatter; distant speech · voices: S1 M, S2 M, S3 F, S4 M, S5 F
EOT 15 yields / 0 holds · INT 2 interruptions / 8 negatives (6 backchannels, 2 laughter) · 2 failed claims (excluded)



Transcript of test_00138

start-end (s) speaker [ground-truth label] text > outcome

0.3–11.7 **S1** [turn] Okay, so this game's called Cascadia, and we're building tile maps kind of like a puzzle. You get to place terrain and animal tokens each turn, and the goal is to connect habitats and get points for animal arrangements. > yields to S5

6.1–6.8 **S4** [backchannel] Uh-huh.

11.7–15.1 **S5** [turn] So, like, do we all share the same board, or does everyone have their own? > yields to S1

15.0–20.5 **S1** [turn] Everybody gets their own board, but the pool of tiles and tokens in the middle is shared. You pick from those on your turn— > cut off

19.6–24.3 **S3** [interruption] Wait, do you have to take both a tile and an animal, or could you, like, just choose one? > takes floor from S1

24.4–31.5 **S3** [turn] Because I remember with Azul there's a turn where you can only pick tiles, so I just wanted to check if it's the same sort of thing. > yields to S1

29.2–29.7 **S2** [backchannel] mm-hm

31.4–38.7 **S1** [turn] Good question! No, in Cascadia you always pick a pair—a tile and an animal token together, and there are four sets to choose from each round. > yields to S2

33.0–34.0 **S2** [backchannel] Ah, okay.

39.3–44.6 **S2** [turn] So, if the animal on the token doesn't match the tile, are you just stuck, or can you swap it or something? > yields to S5

44.7–47.5 **S5** [turn] Yeah, that confused me last time we watched that playthrough. > yields to S1

47.4–57.7 **S1** [turn] You're allowed to place the animal token on any empty terrain with its matching icon. If you can't, you have to discard the animal. But there are nature tokens you earn that can help you mix up your picks if, uh—

54.0–55.2 **S4** [failed claim] Wait a second—

57.8–65.4 **S1** [turn] —if none of the pairs work for you, or you want a better combination, you can spend a nature token to take any tile and animal, not just the paired ones. > yields to S3

63.1–64.9 **S4** [failed claim] Right, right. That makes sense.

65.7–69.6 **S3** [turn] So scoring—uh, is it based on the animals or the terrains, or both? > cut off

69.0–72.4 **S5** [interruption] Like, do we get more points for bigger groups or for having every animal? > takes floor from S3

72.5–75.3 **S5** [turn] I'm just trying to get my strategy right before we start. > yields to S1

73.2–73.8 **S2** [laughter] haha

75.3–87.1 **S1** [turn] Both, actually! You get points for the largest contiguous region of each terrain type, and then points for animal patterns. Each animal has a different scoring card, so like, hawks want to be separated, salmon want to be in chains. > yields to S3

87.0–88.9 **S3** [turn] Oh, so it changes every game a little? > yields to S1

88.8–91.2 **S1** [turn] Exactly. That keeps it from getting stale. > yields to S4

90.0–90.8 **S2** [backchannel] Ah, cool.

90.0–90.8 **S5** [backchannel] Got it.

91.6–95.8 **S4** [turn] So when it's my turn—wait, do I have to finish placing before the next person goes? > yields to S1

95.9–101.4 **S1** [turn] Yeah, we, uh, go one at a time, just so there's no mistakes with the shared pool. But turns are usually pretty fast. > yields to S3

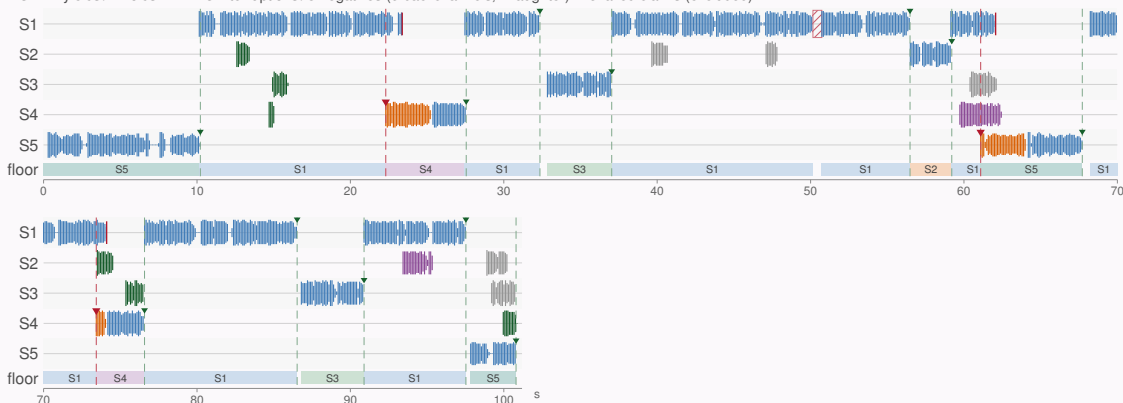
101.3–104.8 **S3** [turn] Alright, I'm ready. Let's see how much I mess up the first round! > yields to S5

104.1–104.6 **S4** [laughter] haha

104.8–106.9 **S5** [turn] Promise we'll go easy for the first game. > yields, end

105.6–106.2 **S2** [backchannel] Yup.

test split · tone: professional and business-like · noise: school hallway, 32.9 dB SNR · voices: S1 M, S2 F, S3 F, S4 M, S5 M
EOT 12 yields / 1 holds · INT 3 interruptions / 8 negatives (6 backchannels, 2 laughter) · 5 failed claims (excluded)



Transcript of test_00175

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–10.2 **S5** [turn] Alright everyone, before we get started, I just want to confirm we're all here for the remote access walkthrough, right? Mark, you'll be leading us through the process? ▷ yields to S1

10.2–18.6 **S1** [turn] Uh, yes, that's right. I, um, I'll go step by step from installing the software to, uh—making sure it's running. If you have any questions at any point, just let me know.

12.6–13.4 **S2** [backchannel] Let's do this!

14.7–15.1 **S4** [backchannel] Cool.

14.9–15.9 **S3** [backchannel] Sounds good.

18.6–23.4 **S1** [turn] Okay, so—um, first, you'll want to download SecureConnect from the company portal. It's the blue icon near the top right. ▷ cut off

22.3–25.3 **S4** [interruption] Wait, which portal? The main dashboard or, like, the IT one? ▷ takes floor from S1

25.4–27.6 **S4** [turn] Sorry, just wanna make sure I'm at the right place. ▷ yields to S1

27.5–32.4 **S1** [turn] Oh, right, um, sorry—I mean the IT support portal, not the general dashboard. Good catch. ▷ yields to S3

32.8–37.0 **S3** [turn] Is that the same one with the HR links? Because I always get those mixed up. ▷ yields to S1

37.1–43.8 **S1** [turn] No, it's a different portal. The IT one has a white background with a blue bar. If you need the link, I can share it after the call.

39.7–40.7 **S2** [failed claim] I see it now, thanks.

43.9–50.2 **S1** [turn] Once you've... downloaded it, just double-click the installer. It should basically walk you through—uh, the steps automatically. ▷ holds, pauses

47.1–47.9 **S2** [failed claim] Wait, wait—

50.7–56.5 **S1** [turn] —yeah, it'll ask for administrator permissions next. You'll need to click yes there or it won't finish installing. ▷ yields to S2

56.5–59.2 **S2** [turn] Oh, okay, so I should just, like, trust it and hit yes? ▷ yields to S1

59.1–62.1 **S1** [turn] Yes, that's totally safe. It's our approved software. ▷ cut off

59.7–62.5 **S4** [laughter] Haha, as long as it's not another Zoom update, I'm good.

60.4–62.1 **S3** [failed claim] That reassures me, thanks.

61.1–64.0 **S5** [interruption] Mark, could you clarify what happens after it's installed— ▷ takes floor from S1

64.2–67.7 **S5** [turn] like, do we log in with the same credentials as our work email? ▷ yields to S1

68.2–74.1 **S1** [turn] Yeah, you use your standard work email and password. Once you log in, you should see a green check mark if the connection's successful. Um, if it's red, it means there's a problem with the network. ▷ cut off

73.4–74.1 **S4** [interruption] What if, uh— ▷ takes floor from S1

73.5–74.6 **S2** [backchannel] That's super clear.

74.2–76.6 **S4** [turn] What if it stays red even after a couple tries? ▷ yields to S1

75.4–76.6 **S3** [backchannel] Awesome, thanks.

76.6–86.5 **S1** [turn] Good question. If it's still red, try restarting your computer first. If that doesn't work, check your Wi-Fi. Sometimes the firewall blocks it, so you might need to contact IT if it still doesn't connect. ▷ yields to S3

86.8–90.9 **S3** [turn] Could you show what the firewall error looks like? Just so I know what to look for. ▷ yields to S1

90.9–97.5 **S1** [turn] Sure, it usually gives a pop-up in the lower right corner—like a little shield icon. If you see that, make sure to click 'allow'. ▷ yields to S5

93.4–95.4 **S2** [laughter] I always forget the restart trick, ha.

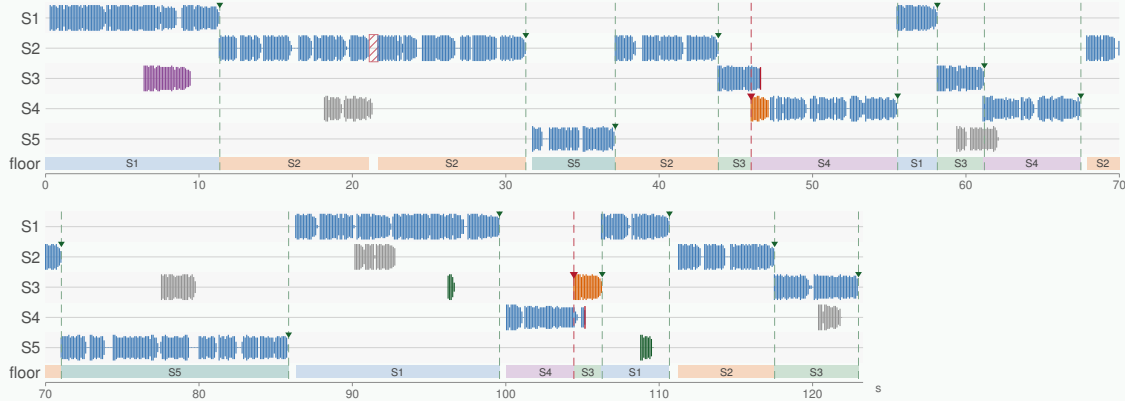
97.8–100.8 **S5** [turn] We're almost at time. Any last questions for Mark? ▷ yields, end

98.9–100.3 **S2** [failed claim] All good here, thanks Mark!

99.2–100.8 **S3** [failed claim] Nope, clear as day for me.

100.0–100.9 **S4** [backchannel] Yeah, thanks man.

dev split · tone: curious and inquisitive · noise: studio room tone, 21.8 dB SNR; keyboard clatter · voices: S1 F, S2 M, S3 M, S4 M, S5 F
EOT 15 yields / 1 holds · INT 2 interruptions / 3 negatives (2 backchannels, 1 laughter) · 5 failed claims (excluded)

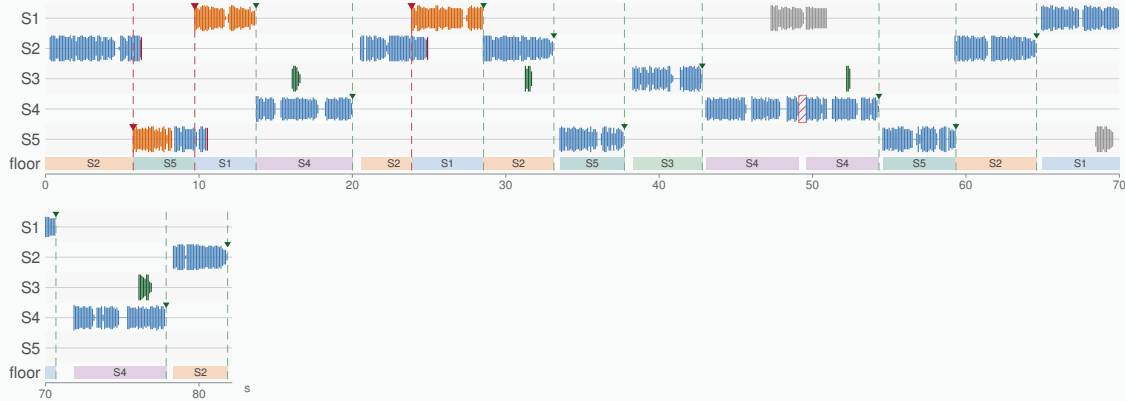


Transcript of dev_00007

start–end (s) speaker [ground-truth label] text ▷ outcome

- 0.3–11.4 **S1** [turn] Okay, so before we start, let's just set some ground rules: book or movie, doesn't matter, but it has to be something you've actually finished recently. No half-watched Netflix series, alright? ▷ yields to S2
- 6.4–9.5 **S3** [laughter] Haha, fair, I definitely fall into that trap sometimes.
- 11.3–21.1 **S2** [turn] So, I'll just jump in—I actually just finished "The Midnight Library" by Matt Haig. It's got this, um, sort of multiverse concept, kind of reminds me of— ▷ holds, pauses
- 18.2–21.3 **S4** [failed claim] Wait, you mean it's like "Everything Everywhere All at Once"—?
- 21.7–31.3 **S2** [turn] —yeah, but more introspective, less chaos. It's about regrets and possibilities. Honestly, I thought it was a bit on the nose, but some chapters were moving. ▷ yields to S5
- 31.7–37.1 **S5** [turn] Interesting. Did you relate to the main character at all, or was it more just the concept that hooked you? ▷ yields to S2
- 37.1–43.9 **S2** [turn] Mostly the concept. It got me thinking about all the tiny choices that lead you somewhere totally different. ▷ yields to S3
- 43.8–46.6 **S3** [turn] That actually sounds right up my alley, I love that kind of thing. ▷ cut off
- 46.0–47.2 **S4** [interruption] Okay but— ▷ takes floor from S3
- 47.3–55.5 **S4** [turn] —my pick totally flips that introspective vibe. I watched "Prey" recently, the new Predator film? It's all action, actually super smart for a blockbuster. ▷ yields to S1
- 55.5–58.1 **S1** [turn] Oh, that's the one set in the 1700s, right? ▷ yields to S3
- 58.1–61.2 **S3** [turn] Right, I heard it's surprisingly thoughtful about indigenous culture. ▷ yields to S4
- 59.4–62.1 **S5** [failed claim] Exactly. It's been getting lots of praise for that.
- 61.1–67.5 **S4** [turn] Yeah, and the action scenes are actually tense, not just loud. It's a nice reset if you're burned out on superhero stuff. ▷ yields to S2
- 67.9–71.0 **S2** [turn] Hmm, I haven't seen a Predator movie since the original. ▷ yields to S5
- 71.0–81.2 **S5** [turn] So, I'll recommend "The Guncle"—a book this time. It's by Steven Rowley and is way funnier than I expected, but it also sneaks up on you with the heart. Really great for summer.
- 77.6–79.8 **S3** [failed claim] Is it more comedy or—is it dramatic too?
- 81.3–85.8 **S5** [turn] —it's both, honestly. There's family drama, but mostly it's just warm and witty. ▷ yields to S1
- 86.3–92.5 **S1** [turn] I'll add that to my list. Anyone read "Sea of Tranquility" yet? It's Emily St. John Mandel's new one—
- 90.2–92.9 **S2** [failed claim] Oh yes, isn't that the time travel one?
- 92.6–99.6 **S1** [turn] —sort of, yeah. Different timelines, but much quieter than, like, a sci-fi epic. It's more about connections between people. ▷ yields to S4
- 96.2–96.7 **S3** [backchannel] mm-hm
- 100.0–105.2 **S4** [turn] Look, I gotta ask, does anyone here actually prefer movies to books for this kind of story? Because— ▷ cut off
- 104.4–106.3 **S3** [interruption] No way, books always win for me. ▷ takes floor from S4 ▷ yields to S1
- 106.2–110.7 **S1** [turn] Honestly, depends on my mood. But books let you get inside someone's head more. ▷ yields to S2
- 108.8–109.6 **S5** [backchannel] Absolutely.
- 111.3–117.5 **S2** [turn] I think certain concepts just work better in print, though film is unbeatable for pacing and visuals. ▷ yields to S3
- 117.5–123.0 **S3** [turn] Haha, honestly, that's the ideal—read the book, then watch the movie adaptation and argue about which is better. ▷ yields, end
- 120.4–121.9 **S4** [failed claim] Can't argue with that.

dev split · tone: casual and friendly · noise: office hum, 27.3 dB SNR · voices: S1 F, S2 F, S3 M, S4 F, S5 F
EOT 12 yields / 1 holds · INT 3 interruptions / 4 negatives (4 backchannels, 0 laughter) · 2 failed claims (excluded)

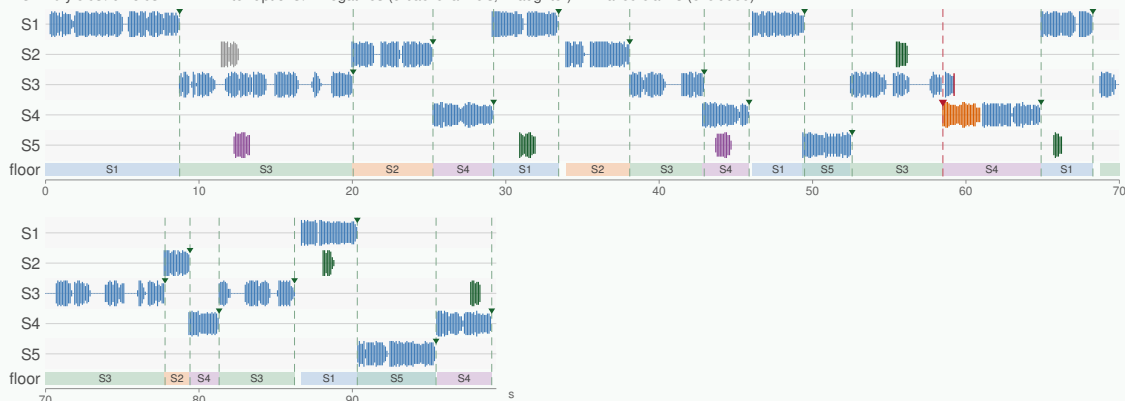


Transcript of dev_00016

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.2 **S2** [turn] Alright, so we're looking to get Dr. Mendes's calendar filled for next week. First, does anyone have a sense if she's open on Tuesday mornings? ▷ cut off
5.7–8.3 **S5** [interruption] I think she usually has that research block, right? ▷ takes floor from S2
8.4–10.5 **S5** [turn] Or at least, last week she did? Unless it changed— ▷ cut off
9.7–13.7 **S1** [interruption] —Are we sure that wasn't canceled this month? She mentioned her schedule was shifting. ▷ takes floor from S5 ▷ yields to S4
13.7–20.0 **S4** [turn] That's what I heard as well. She moved the research block to Wednesdays for June. So Tuesday's probably open. ▷ yields to S2
16.1–16.6 **S3** [backchannel] Right.
20.5–24.9 **S2** [turn] Great, that makes things easier. So should I start with morning slots or afternoons? ▷ cut off
23.9–28.6 **S1** [interruption] Wait, how many patients do we actually need to fit in on Tuesday? Because if we're overbooked— ▷ takes floor from S2 ▷ yields to S2
28.5–33.1 **S2** [turn] —Five, max. The three regulars and then two follow-ups from last week's reschedules. ▷ yields to S5
31.3–31.7 **S3** [backchannel] Uh-huh.
33.5–37.7 **S5** [turn] Morning's usually better for Mrs. Hsu, she likes those first appointments. ▷ yields to S3
38.3–42.8 **S3** [turn] But Mr. Peters can't do before ten. He told me again yesterday. ▷ yields to S4
43.1–49.1 **S4** [turn] So if we put Mrs. Hsu in at eight forty-five and Mr. Peters at ten fifteen, that leaves a gap— ▷ holds, pauses
47.3–51.0 **S1** [failed claim] —Doesn't that conflict with the staff meeting, or am I just making that up?
49.6–54.3 **S4** [turn] —no, that's a valid point. The staff meeting is at nine thirty. We'd need to buffer that. ▷ yields to S5
52.2–52.5 **S3** [backchannel] Mm.
54.6–59.3 **S5** [turn] Could we squeeze one in right after the meeting, like ten o'clock on the dot, or is that pushing it? ▷ yields to S2
59.3–64.6 **S2** [turn] If everyone's on time, it's doable. But Dr. Mendes tends to run late out of meetings. ▷ yields to S1
64.9–70.7 **S1** [turn] Maybe we just start patient slots at ten fifteen, stack the rest back-to-back, and warn them appointments may run tight? ▷ yields to S4
68.5–69.6 **S5** [failed claim] That works for me.
71.9–77.9 **S4** [turn] I'll block the calendar now and send appointment reminders. Let me know if any conflicts pop up last minute. ▷ yields to S2
76.1–76.9 **S3** [backchannel] Sounds good.
78.3–81.9 **S2** [turn] Thanks everyone. I'll email confirmations and we'll see how it goes. ▷ yields, end

test split · tone: curious and inquisitive · noise: quiet apartment, 30.4 dB SNR; alarm beep · voices: S1 M, S2 M, S3 F, S4 M, S5 F
EOT 19 yields / 0 holds · INT 1 interruptions / 7 negatives (5 backchannels, 2 laughter) · 1 failed claims (excluded)

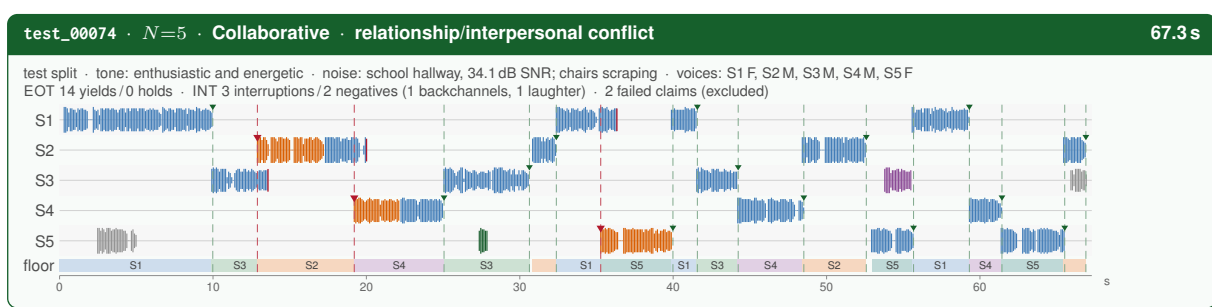


Transcript of test_00052

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.7 **S1** [turn] So um, the question is, how are we getting everyone to Anna's piano recital on Saturday? Because I think we've got, like, three different places people are coming from— ▷ yields to S3
8.7–13.5 **S3** [turn] It's only a logistical nightmare if you think about it too hard.
11.5–12.6 **S2** [failed claim] I mean, if we just—
12.3–13.4 **S5** [laughter] Haha, right.
13.6–20.1 **S3** [turn] —but seriously, what time does it start again? Four? Or was that last time? ▷ yields to S2
20.0–25.3 **S2** [turn] No, it's four this time too. But we need to be there at like, three forty-five, according to the email. ▷ yields to S4

25.3–29.2 **S4** [turn] Do we need to factor in parking again? Because last time it was a total zoo. > yields to S1
 29.1–33.4 **S1** [turn] Yeah, uh, I think it'll be worse, actually? There's some event at the community center too. > yields to S2
 30.9–32.0 **S5** [backchannel] Oh, really?
 33.9–38.1 **S2** [turn] Maybe we carpool, then? I can grab Mom and Alex on the way if that helps. > yields to S3
 38.1–42.9 **S3** [turn] Assuming Alex is ready on time, which is a bold assumption. > yields to S4
 42.8–45.9 **S4** [turn] Hey, I'm totally punctual, at least on weekends. Mostly. > yields to S1
 43.7–44.7 **S5** [laughter] Haha, mm-hm.
 46.1–49.5 **S1** [turn] Do we need snacks for after, or is it just the post-recital cookies? > yields to S5
 49.4–52.6 **S5** [turn] I think they're only doing juice boxes this time, actually. > yields to S3
 52.5–59.2 **S3** [turn] I can bring some of those energy bars just in case. Someone's always starving. > cut off
 55.5–56.3 **S2** [backchannel] Good idea.
 58.5–61.0 **S4** [interruption] So, logistically, if you pick up Mom and Alex— > takes floor from S2
 61.1–64.9 **S4** [turn] —that just leaves me to get Jamie, right? Because I'm coming from work earlier. > yields to S1
 64.9–68.3 **S1** [turn] Wait, Jamie has soccer till three, so that's... a little tight, isn't it? > yields to S3
 65.7–66.3 **S5** [backchannel] Right.
 68.7–77.8 **S3** [turn] We could switch it up. I can get Jamie and bring her with me. I drive fast. Well, legally fast. > yields to S2
 77.7–79.4 **S2** [turn] That actually makes things simpler. > yields to S4
 79.3–81.3 **S4** [turn] Oh, but then you have to detour all the way south. > yields to S3
 81.3–86.2 **S3** [turn] It's fine. I'll grab coffee on the way, make it an adventure. > yields to S1
 86.7–90.3 **S1** [turn] Okay—so to recap, Alex and Mom with you, Greg, and Jamie with Riley? > yields to S5
 88.1–88.8 **S2** [backchannel] That's right.
 90.4–95.5 **S5** [turn] Sounds like it'll all work out. Uh, as long as we remember to actually leave on time. > yields to S4
 95.5–99.1 **S4** [turn] Yeah, I'll set like, three timers. Or, you know, just follow the scent of panic. > yields, end
 97.7–98.4 **S3** [backchannel] Classic.

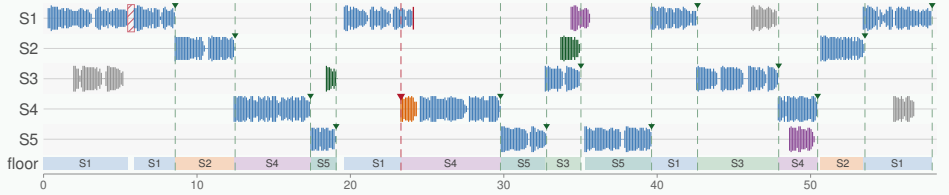


Transcript of test_00074

start–end (s) speaker [ground-truth label] text > outcome

0.3–6.6 **S1** [turn] Alright, so I'm just gonna say it. If we keep dancing around what bothered Sam last night, nobody's gonna actually fix anything.
 2.5–5.0 **S5** [failed claim] Wait, wait, I actually love this energy! But—
 6.7–10.0 **S1** [turn] —I mean, come on, Sam literally left the table. That doesn't just happen. > yields to S3
 9.9–13.6 **S3** [turn] Right, that was so awkward. I almost spilled my drink because I was staring after him like, what? > cut off
 12.9–17.2 **S2** [interruption] I think, um, before we pile on Sam, maybe it's helpful if we actually ask him— > takes floor from S3
 17.3–20.0 **S2** [turn] —well, ask him directly, like, what was going on? Otherwise we're guessing, right? > cut off
 19.2–22.2 **S4** [interruption] Yeah, but, I mean, we all saw the look when Jamie mentioned the trip— > takes floor from S2
 22.3–25.1 **S4** [turn] He got super quiet. Maybe it's about feeling left out? > yields to S3
 25.1–28.2 **S3** [turn] Oh, totally! He did that thing with his hands, you know, where he like—
 27.4–27.9 **S5** [backchannel] Fidgets!
 28.2–30.6 **S3** [turn] —yeah, exactly. It's like, textbook Sam. > yields to S2
 30.8–32.4 **S2** [turn] So, should we just text him, or— > yields to S1
 32.4–36.3 **S1** [turn] No, I think a call is better. Messages get misread. If someone volunteers, I can— > cut off
 35.3–40.0 **S5** [interruption] Ohmygod I'll do it! But only if at least two of you promise to back me up if he gets weird. > takes floor from S1 > yields to S1
 39.9–41.6 **S1** [turn] Deal, seriously, I'm right there. > yields to S3
 41.5–44.2 **S3** [turn] Yeah, count me in. Solidarity squad, activate! > yields to S4
 44.2–48.5 **S4** [turn] Haha, okay, but what do we actually say? Like, uh, 'hey, you stormed off, spill?' Or... > yields to S2
 48.4–52.6 **S2** [turn] I think gentle's better. Maybe ask if he felt left out or if something else was up. > yields to S5
 52.9–55.7 **S5** [turn] 'Hey Sam, are you mad at us?' No, okay, too much. > yields to S1
 53.8–55.6 **S3** [laughter] Haha, please don't lead with that.
 55.6–59.3 **S1** [turn] Just say we missed him, and if there's something bugging him, we actually wanna know. > yields to S4
 59.3–61.4 **S4** [turn] Right, honest but not, like, cornering him. > yields to S5
 61.4–65.5 **S5** [turn] Okay, I'm dialing. Put me on speaker, and no one say anything weird—yeah? > yields to S2
 65.5–66.9 **S2** [turn] Promise. Quiet as a mouse. > yields, end
 65.9–67.0 **S3** [failed claim] We'll see about that!

test split · tone: enthusiastic and energetic · noise: living room tv distant, 33.4 dB SNR · voices: S1 M, S2 M, S3 F, S4 F, S5 F
EOT 13 yields / 1 holds · INT 1 interruptions / 4 negatives (2 backchannels, 2 laughter) · 3 failed claims (excluded)

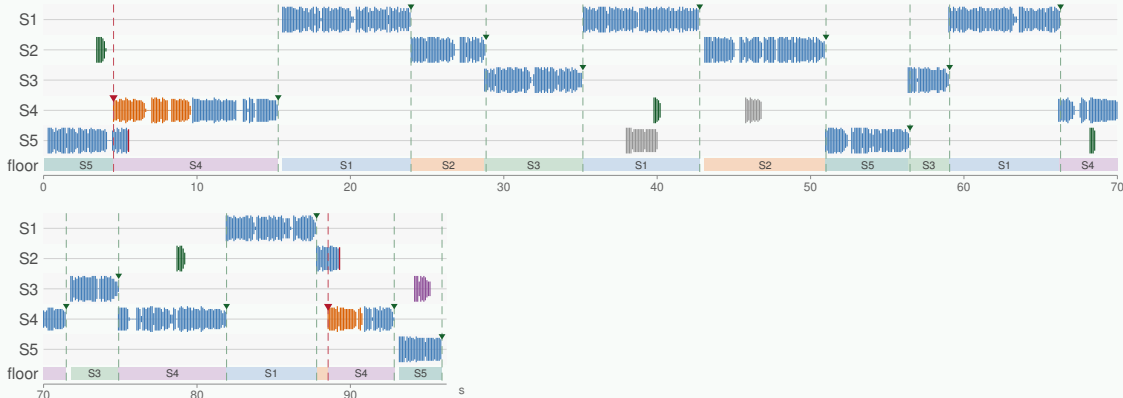


Transcript of test_00099

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–5.5 **S1** [turn] Okay, so are we doing Codenames or Betrayal first? I've honestly been hyped for Betrayal all week— ▷ holds, pauses
2.0–5.2 **S3** [failed claim] Wait, Codenames is way quicker, we can get into the groove—
5.9–8.6 **S1** [turn] —but Codenames would be a good warm-up, yeah, you're right. ▷ yields to S2
8.6–12.5 **S2** [turn] I uh, I vote for Codenames first too. Just so there's... less setup at the start. ▷ yields to S4
12.4–17.4 **S4** [turn] Right, and I brought snacks, so if we play something fast, we can snack in between games. ▷ yields to S5
17.4–19.1 **S5** [turn] Oh, that's such a good idea. ▷ yields to S1
18.4–19.1 **S3** [backchannel] Exactly.
19.6–24.1 **S1** [turn] Alright, so Codenames it is. Who wants to be spymasters? I always mess up the clues— ▷ cut off
23.3–24.4 **S4** [interruption] I'll do it if— ▷ takes floor from S1
24.6–29.8 **S4** [turn] I'll do it if someone actually wants to guess on my team. I'm not going up against Maya alone. ▷ yields to S5
29.8–32.8 **S5** [turn] Haha, no worries, I'll be on your team. Maya's ruthless. ▷ yields to S3
32.7–35.0 **S3** [turn] I'm not ruthless, I just... play to win, okay? ▷ yields to S5
33.7–35.0 **S2** [backchannel] Yeah, seriously.
34.4–35.6 **S1** [laughter] Haha, proven facts.
35.3–39.6 **S5** [turn] But I think we should mix up teams from last time. Tim and I always end up together. ▷ yields to S1
39.6–42.6 **S1** [turn] Oh, oops, my bad! Uh—wait, what if I'm spymaster for once? ▷ yields to S3
42.6–47.1 **S3** [turn] You always overthink it though, last time you said 'nature, one' and meant broccoli.
46.2–47.8 **S1** [failed claim] That was perfectly logical—
47.2–47.9 **S3** [turn] —for you maybe! ▷ yields to S4
47.9–50.5 **S4** [turn] He does have broccoli energy though, I have to say. ▷ yields to S2
48.6–50.2 **S5** [laughter] Haha, gentle but strong.
50.7–53.5 **S2** [turn] I honestly just want to see what clue he'd give for 'penguin'. ▷ yields to S1
53.4–57.9 **S1** [turn] Oh, challenge accepted! Just wait, you're all going to regret letting me have the clue cards. ▷ yields, end
55.4–56.8 **S4** [failed claim] I can't wait for this.

test split · tone: professional and business-like · noise: quiet office, 35.0 dB SNR; distant speech · voices: S1 F, S2 M, S3 M, S4 F, S5 F
EOT 15 yields / 0 holds · INT 2 interruptions / 5 negatives (4 backchannels, 1 laughter) · 2 failed claims (excluded)



Transcript of test_00103

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–5.5 **S5** [turn] Alright, so should we go with Settlers of Catan or try something new tonight? We've done Catan like three times this month, but I'm honestly fine either way. ▷ cut off
3.5–4.1 **S2** [backchannel] I think—
4.6–9.6 **S4** [interruption] Oooh, what if we finally tried Wingspan? I saw it in the cabinet and it looked so colorful. ▷ takes floor from S5
9.7–15.3 **S4** [turn] I've read the rules and it's not that complicated, I promise. And, uh, there are birds involved, so... ▷ yields to S1
15.6–24.0 **S1** [turn] Haha, birds as a selling point, love that. But I'm game for trying something new. I've actually heard Wingspan is strategic, just different from Catan. ▷ yields to S2
24.0–28.8 **S2** [turn] I'm definitely open, though I'd need a quick rules walkthrough. I haven't played that one before. ▷ yields to S3
28.8–35.2 **S3** [turn] Um, is it, uh, really that much harder than Catan? 'Cause I still get lost with, you know, trading sometimes. ▷ yields to S1
35.2–38.8 **S1** [turn] No, I don't think it's harder, just, uh, different types of decisions.
38.0–40.1 **S5** [failed claim] Yeah, actually, the rounds are simpler, I think—
38.9–42.8 **S1** [turn] —you get to focus more on your own objectives, less on blocking people. ▷ yields to S2
39.8–40.2 **S4** [backchannel] right
43.1–46.9 **S2** [turn] So, uh, what's the win condition in Wingspan? Most points, or something else?
45.8–46.9 **S4** [failed claim] It's points, but you—
47.0–51.0 **S2** [turn] —get a scorepad, and there's, you know, eggs and birds and habitat bonuses. ▷ yields to S5

51.0–56.5 **S5** [turn] I love that there are eggs. It sounds, like, way less cutthroat than stealing longest road all the time. ▷ yields to S3

56.4–59.1 **S3** [turn] Haha, yeah. Uh, I mean, I always lose longest road. ▷ yields to S1

59.0–66.3 **S1** [turn] Oh, same here. But, honestly, with Wingspan, it's more individual tableau-building. No one can really 'steal' your progress, right? ▷ yields to S4

66.2–71.5 **S4** [turn] Yup, exactly! Plus, you get to make your own little bird sanctuary, and the art is so good. ▷ yields to S3

68.2–68.6 **S5** [backchannel] right

71.8–74.9 **S3** [turn] So, can someone walk me through setup? Like, just the basics? ▷ yields to S4

74.9–81.9 **S4** [turn] Yeah, I'll do it! Uh, first we deal out five bird cards, then—you get two bonus cards, eggs, and some food tokens. ▷ yields to S1

78.7–79.3 **S2** [backchannel] Uh-huh.

81.9–87.8 **S1** [turn] And every turn you choose one action—draw cards, lay eggs, or gain food. That's pretty much it at first. ▷ yields to S2

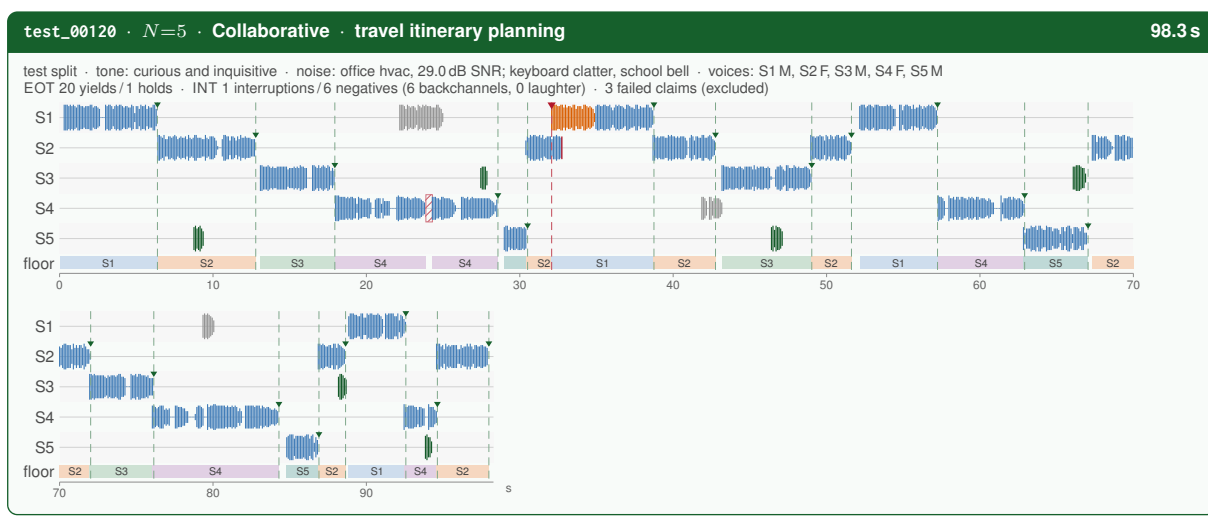
87.8–89.3 **S2** [turn] Got it, that seems simple enough. ▷ cut off

88.5–90.8 **S4** [interruption] Do you want me to shuffle while you explain, or— ▷ takes floor from S2

90.9–92.9 **S4** [turn] Actually, I'll just shuffle either way! ▷ yields to S5

93.2–96.0 **S5** [turn] Perfect, let's get the snacks over here too while we set up. ▷ yields, end

94.2–95.2 **S3** [laughter] Haha, priorities.



Transcript of test_00120

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–6.4 **S1** [turn] Uh, okay, should we start by picking the first city? Like, um, maybe Rome because it's got, you know, all the history and stuff. ▷ yields to S2

6.4–12.8 **S2** [turn] I really think Florence would actually make more sense to start with since it's less chaotic. Rome gets packed, especially in June. ▷ yields to S3

8.8–9.4 **S5** [backchannel] Yeah, true.

13.1–17.9 **S3** [turn] But if we land in Rome, isn't it, um, easier to just start there? I'm just thinking about the flights. ▷ yields to S4

18.0–23.9 **S4** [turn] Except the train from Rome to Florence is, like, super fast. We could land, get straight on a train— ▷ holds, pauses

22.2–25.0 **S1** [failed claim] But wait, wouldn't we be, uh, too tired after the flight—

24.3–28.6 **S4** [turn] —and then just nap on the train instead of wasting a day in Rome all jetlagged? ▷ yields to S5

27.5–27.9 **S3** [backchannel] Yeah

29.0–30.5 **S5** [turn] That actually sounds kind of smart. ▷ yields to S2

30.4–32.7 **S2** [turn] See, and then by the time we get to Florence, it's afternoon, we check in, we eat— ▷ cut off

32.1–34.9 **S1** [interruption] But the places might not let us check in that early, though? ▷ takes floor from S2

35.0–38.7 **S1** [turn] I mean, last time I went, check-in was, uh, like three PM everywhere. ▷ yields to S2

38.7–42.8 **S2** [turn] Right, but we can just leave our bags at the hotel. Most places let you do that. ▷ yields to S3

41.9–43.2 **S4** [failed claim] Yeah, it's pretty normal.

43.2–49.0 **S3** [turn] Also, if we start in Florence, we can do Tuscany early in the trip, which, um, honestly sounds less exhausting? ▷ yields to S2

46.4–47.2 **S5** [backchannel] Yeah, totally.

49.0–51.6 **S2** [turn] Okay, so, Florence first, then? For, what, three nights? ▷ yields to S1

52.2–57.2 **S1** [turn] I guess, but, um, can we maybe do four nights? Just because the Uffizi takes, like, a whole day. ▷ yields to S4

57.3–62.9 **S4** [turn] I mean, I think the city itself is small enough that three is fine, but four's not unreasonable. ▷ yields to S5

62.8–67.0 **S5** [turn] If we do day trips out to Siena or Pisa, then, um, four makes sense. ▷ yields to S2

66.1–66.9 **S3** [backchannel] Exactly.

67.3–72.0 **S2** [turn] Alright, four for Florence. What about Rome after, or do we swing by the coast, like Cinque Terre, first? ▷ yields to S3

72.0–76.1 **S3** [turn] Um, my only worry with Cinque Terre is the crowds. It, it gets packed on weekends. ▷ yields to S4

76.0–82.0 **S4** [turn] But if we hit it midweek, it could be fine. Plus, it's kind of on the way if we're heading south after.

79.4–80.1 **S1** [failed claim] Yeah but—

82.1–84.3 **S4** [turn] —we just need to check the train schedules, though. ▷ yields to S5

84.8–86.9 **S5** [turn] I can look up train times, if that's helpful. ▷ yields to S2

86.9–88.6 **S2** [turn] Yeah, that'd be great. Thanks, Ben. ▷ yields to S1

88.2–88.8 **S3** [backchannel] Nice.

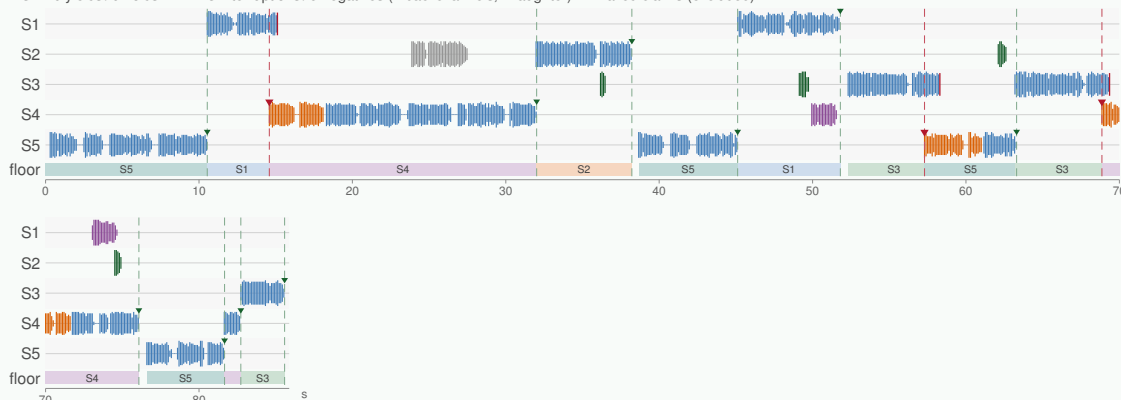
88.9–92.6 **S1** [turn] So, uh, maybe Florence, Cinque Terre, then Rome? Or should we, um, swap that? ▷ yields to S4

92.5–94.6 **S4** [turn] I think that's the most efficient, actually. ▷ yields to S2

93.8–94.3 **S5** [backchannel] Mm-hm.

94.6–98.0 **S2** [turn] Alright, let's start drafting a rough schedule? I can get a spreadsheet going. ▷ yields, end

test split · tone: thoughtful and contemplative · noise: office hum, 23.7 dB SNR; paper shuffle · voices: S1 M, S2 M, S3 M, S4 F, S5 M
EOT 10 yields / 0 holds · INT 3 interruptions / 6 negatives (4 backchannels, 2 laughter) · 1 failed claims (excluded)

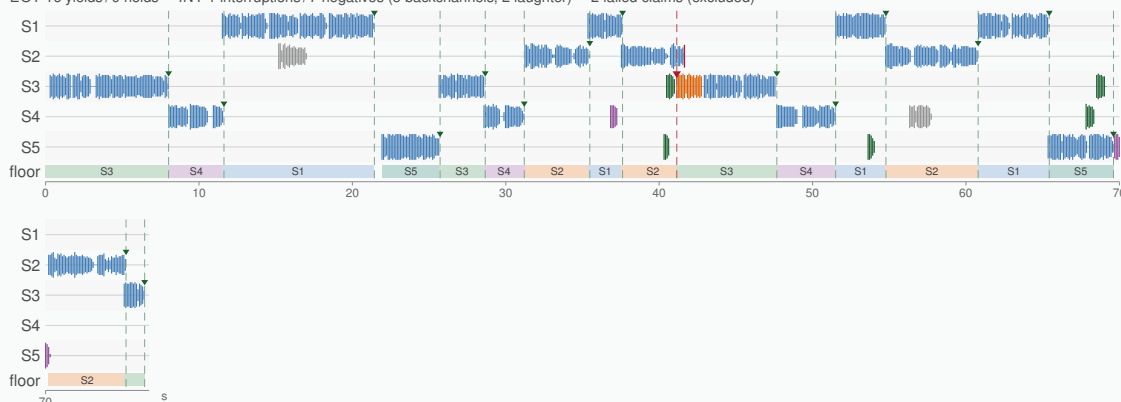


Transcript of test_00148

start-end (s) speaker [ground-truth label] text > outcome

0.3–10.5 **S5** [turn] All right, welcome back everyone, let's dive right in—tonight's game really turned in the fourth quarter, but was that more about the offense or the defense stepping up? > yields to S1
10.5–15.1 **S1** [turn] I'd say the defense, honestly. You look at how they kept forcing turnovers, it reminded me of, uh, playoff intensity. Like, that desperation— > cut off
14.6–18.2 **S4** [interruption] Ooh but did you see the bench energy? I think that's what swung things— > takes floor from S1
18.3–27.4 **S4** [turn] I mean, you had reserves just hopping up, waving towels, hyping the crowd every play. The momentum totally shifted when the big steal happened. That's not—
23.9–27.5 **S2** [failed claim] Um, sorry, but—but don't you think it's actually the point guard's pacing?
27.6–32.0 **S4** [turn] —just hype, it's actual tactical stuff, you know? Defensive rotations, energy. > yields to S2
32.0–38.2 **S2** [turn] No, uh, I mean, the way he controlled the tempo, especially after halftime? That's when they started really pulling ahead. > yields to S5
36.2–36.6 **S3** [backchannel] yeah
38.7–45.1 **S5** [turn] That's a fair point about pacing. But what about the big man? He changed a lot inside, altered shots. > yields to S1
45.1–51.8 **S1** [turn] Exactly. He was camping in the paint, really made them adjust. You could see, um, they hesitated before driving after halftime. > yields to S3
49.1–49.8 **S3** [backchannel] Right, right.
49.9–51.6 **S4** [laughter] Haha, intimidation factor.
52.3–58.3 **S3** [turn] Honestly, I think that third quarter run started with the coach's timeout. Like, he called it right as they were starting to slip, and then— > cut off
57.3–61.1 **S5** [interruption] Oh, do you think that was a momentum thing or just, uh, tactical? > takes floor from S3
61.2–63.3 **S5** [turn] Because the substitutions right after— > yields to S3
62.1–62.7 **S2** [backchannel] mm-hm
63.2–69.4 **S3** [turn] —yeah, that sparked it! The rookie came in, hit a three, and the energy just snapped back. Sometimes a quick sub does more than any speech. > cut off
68.9–71.7 **S4** [interruption] No way, that three was pure luck, did you see the bounce? > takes floor from S3
71.7–76.1 **S4** [turn] Like, it rolled around forever! But hey, sometimes you need a little magic, right? > yields to S5
73.0–74.7 **S1** [laughter] Haha, yeah, ball don't lie.
74.5–74.9 **S2** [backchannel] uh-huh
76.6–81.7 **S5** [turn] So are we saying it's a blend? Coaching, defense, a little luck—it all comes together? > yields to S4
81.6–82.7 **S4** [turn] That's sports, right? > yields to S3
82.7–85.6 **S3** [turn] Exactly. Unpredictable, that's what keeps us watching. > yields, end

test split · tone: enthusiastic and energetic · noise: cafe, 27.9 dB SNR; keyboard clatter; distant speech · voices: S1 F, S2 M, S3 F, S4 F, S5 M
EOT 16 yields / 0 holds · INT 1 interruptions / 7 negatives (5 backchannels, 2 laughter) · 2 failed claims (excluded)



Transcript of test_00183

start-end (s) speaker [ground-truth label] text > outcome

0.3–8.0 **S3** [turn] Okay, so I have to start—I loved the ending! Like, was anyone else totally not expecting her to just leave the letter on the train like that? > yields to S4
8.0–11.6 **S4** [turn] Oh my gosh, yes! I was yelling at my book, like, "What are you doing?" > yields to S1
11.5–16.5 **S1** [turn] Honestly, I thought she was finally going to meet her father. The letter felt so important—
15.2–17.0 **S2** [failed claim] Wait, wait, so you really believed—

16.6–21.4 **S1** [turn] —that she'd actually go through with it? I mean, she was so nervous the whole book, I wasn't sure. > yields to S5

21.9–25.7 **S5** [turn] Um, I-I kind of thought she'd lose it by accident, not like, on purpose. > yields to S3

25.7–28.7 **S3** [turn] Yeah, same! She just, like, set it down so casually. I was shocked. > yields to S4

28.6–31.2 **S4** [turn] Right? And then that guy, what's his name, oh, Ben— > yields to S2

31.2–35.5 **S2** [turn] Ben! Yes, I loved his reaction. He looked so bewildered. You could just picture— > yields to S1

35.4–37.6 **S1** [turn] Oh, totally, he was like, "Wait, are you sure?" > yields to S2

36.8–37.3 **S4** [laughter] haha

37.5–41.6 **S2** [turn] But, I mean, did any of you buy that she'd really forgiven her dad? I just don't get it— > cut off

40.3–40.7 **S5** [backchannel] mm-hm

40.5–41.1 **S3** [backchannel] No way.

41.1–42.9 **S3** [interruption] But maybe that's why she didn't send it. > takes floor from S2

42.9–47.7 **S3** [turn] Like, she wrote this whole emotional confession, but when she actually had the chance, she panicked. > yields to S4

47.7–51.5 **S4** [turn] Yeah! And it's so real, right? I mean, it's scary—putting yourself out there. > yields to S1

51.5–54.8 **S1** [turn] Exactly. I think it's more about her than her dad, honestly. > yields to S2

53.6–54.1 **S5** [backchannel] right

54.8–58.2 **S2** [turn] See, that's what I wanted to talk about! Because I just felt so let down—

56.3–57.8 **S4** [failed claim] But didn't it feel honest, though?

58.3–60.8 **S2** [turn] —by the ending. I wanted catharsis, I guess. > yields to S1

60.8–65.4 **S1** [turn] I get that! But maybe that was the point—it's not all tied up, life's complicated. > yields to S5

65.4–69.6 **S5** [turn] Yeah, I-I liked that it sort of just... lingered. Like, you don't always get closure. > yields to S2

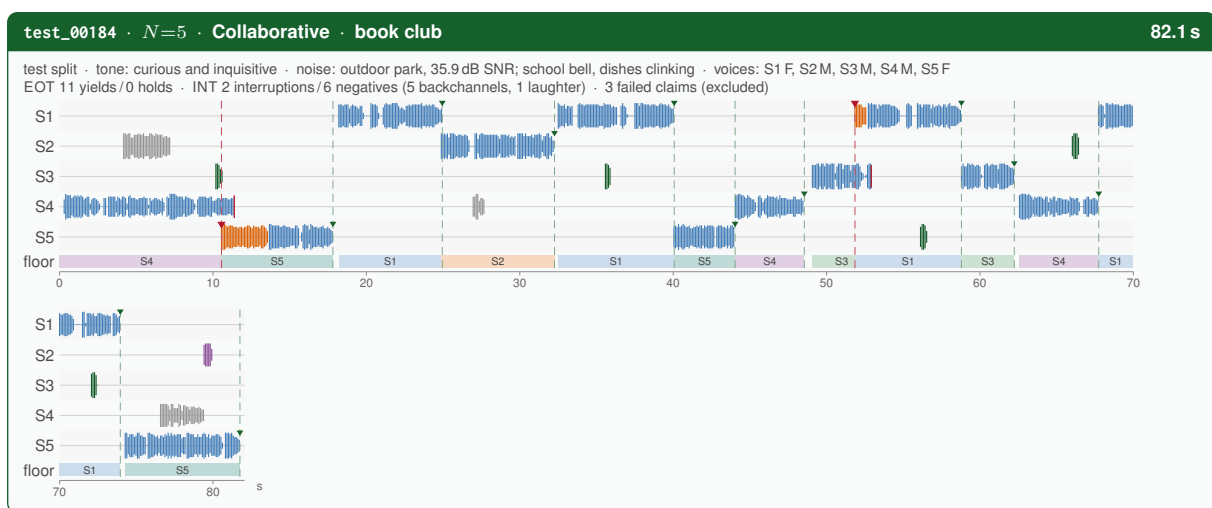
67.8–68.4 **S4** [backchannel] Agreed.

68.5–69.1 **S3** [backchannel] Totally!

69.7–70.3 **S5** [laughter] haha yeah

70.2–75.2 **S2** [turn] Alright. Next time, can we pick something with a happy ending? I need a break from feelings, guys. > yields to S3

75.1–76.5 **S3** [turn] Deal! Who's got suggestions? > yields, end



Transcript of test_00184

start-end (s) speaker [ground-truth label] text > outcome

0.3–6.8 **S4** [turn] So, is it weird that I actually liked the ending? Like, I know most people found it frustrating but I felt it fit the rest of the book.

4.2–7.3 **S2** [failed claim] Wait, um, do you mean that you think it was satisfying the way she just—

7.0–11.4 **S4** [turn] —just walked out? Yeah! That's what I mean. She made her own choice, even if it made no one happy. > cut off

10.2–10.6 **S3** [backchannel] right

10.6–13.6 **S5** [interruption] I actually felt relief when she did that, like she was finally— > takes floor from S4

13.7–17.8 **S5** [turn] —finally allowing herself a little peace, you know? Even if it cost her those relationships. > yields to S1

18.2–25.0 **S1** [turn] But can we really call that peace? I mean, she doesn't resolve anything, she just abandons everything unresolved. > yields to S2

24.9–29.8 **S2** [turn] I, uh, I think that was the point, maybe? Like, the author isn't interested in closure as much as—

27.0–27.8 **S4** [failed claim] —the messiness?

29.9–32.3 **S2** [turn] —yeah, the messiness of real life, exactly. > yields to S1

32.5–40.1 **S1** [turn] I see your point, but there's a difference between messy and incomplete, right? I kept waiting for her to at least call her brother back. > yields to S5

35.6–36.0 **S3** [backchannel] mm-hm

40.0–44.0 **S5** [turn] Do you think the author wants us to feel frustrated by that, or is it just a flaw? > yields to S4

44.1–48.5 **S4** [turn] I took it as intentional? Like, she wants us as uncomfortable as the characters are. > yields to S3

49.1–52.9 **S3** [turn] I, um, I struggled a bit because the pacing threw me off. The last two chapters felt really rushed. > cut off

51.9–52.6 **S1** [interruption] But isn't that— > takes floor from S3

52.7–58.8 **S1** [turn] Isn't that a reflection of her unraveling? Like, things spiral so quickly once she makes that decision. > yields to S3

56.1–56.6 **S5** [backchannel] yes

58.8–62.2 **S3** [turn] Maybe. But I wanted more... um, emotional payoff, I guess. > yields to S4

62.6–67.7 **S4** [turn] Yeah, I get that, but sometimes a book just wants to leave you with a question mark instead of a period. > yields to S1

66.0–66.5 **S2** [backchannel] right

67.7–74.0 **S1** [turn] Still, I would've liked at least a scene between her and her mother. That thread just kind of disappeared, didn't it? > yields to S5

72.1–72.5 **S3** [backchannel] yeah

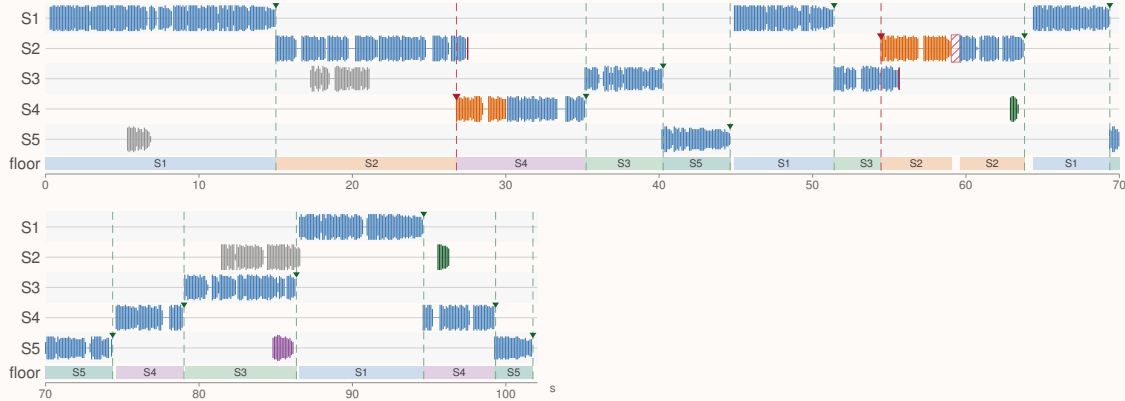
74.3–78.2 **S5** [turn] I was waiting for that too. Maybe in the sequel? She hinted at one in an interview.

76.6–79.4 **S4** [failed claim] Oh, really? Wait, is that official or just a rumor?

78.3–81.8 **S5** [turn] Haha, she said she's writing something 'related'—whatever that means. > yields, end

79.4–80.0 **S2** [laughter] haha

dev split · tone: curious and inquisitive · noise: dorm room, 31.8 dB SNR; notification chime, chairs scraping · voices: S1 F, S2 M, S3 F, S4 F, S5 F
EOT 13 yields / 1 holds · INT 2 interruptions / 3 negatives (2 backchannels, 1 laughter) · 3 failed claims (excluded)

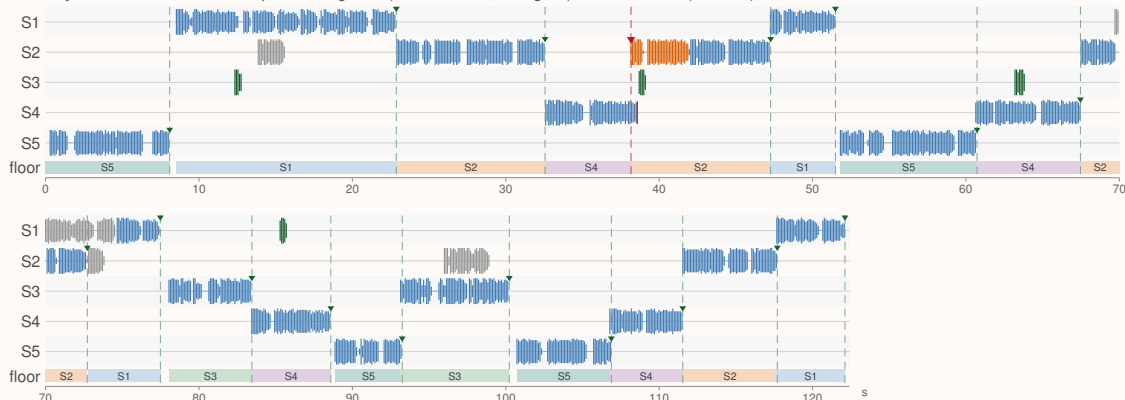


Transcript of dev_00011

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.3 **S1** [turn] Okay, before we go in circles again, can we please clarify which error code actually popped up on the dashboard? Just, let's start there.
5.3–6.9 **S5** [failed claim] Well, that's what I'm saying—
8.4–15.0 **S1** [turn] —because if it's the 503, that's network, but a 504 is timeout, and those need different fixes. ▷ yields to S2
15.0–21.7 **S2** [turn] Sure, but I double-checked, and it was a 504, not a 503. I literally screenshotted it.
17.3–21.2 **S3** [failed claim] Wait, are you positive? Because it was super confusing on my screen, like—
21.8–27.5 **S2** [turn] —the logs also show 504 repeated, and it happened right after we changed the router settings. ▷ cut off
26.8–30.0 **S4** [interruption] Okay, everyone, let's step back. What I'd like to know is— ▷ takes floor from S2
30.1–35.2 **S4** [turn] —has anyone actually tested the connection *after* the server reboot? Or are we just theorizing? ▷ yields to S3
35.2–40.3 **S3** [turn] Um, I did! Right after it came back up, but it still wouldn't load the dashboard, so I'm not sure if— ▷ yields to S5
40.2–44.6 **S5** [turn] Exactly! That's my point, it can't just be a router issue if the reboot didn't fix it. ▷ yields to S1
44.9–51.4 **S1** [turn] Hold on. Are you saying you waited for the full init process before testing? Because that takes at least four minutes. ▷ yields to S3
51.4–55.6 **S3** [turn] I waited like, maybe two? Maybe that's not long enough, but it sounded like it was ready. ▷ cut off
54.5–59.0 **S2** [interruption] It might *sound* ready but the services lag behind— ▷ takes floor from S3 ▷ holds, pauses
59.6–63.8 **S2** [turn] —which is why I said, check the backend logs, not just the web UI. ▷ yields to S1
62.9–63.4 **S4** [backchannel] Right.
64.4–69.4 **S1** [turn] But the logs don't tell us if there's a DNS issue upstream, which would explain the timeout. ▷ yields to S5
69.4–74.4 **S5** [turn] Exactly! So why aren't we focusing on the DNS logs instead? Has anyone actually checked them? ▷ yields to S4
74.6–79.0 **S4** [turn] Because there's no indication that's where the error's coming from, not yet anyway. ▷ yields to S3
79.1–82.4 **S3** [turn] But how would we know if we don't look? I mean, what if it's intermittent?
81.5–86.6 **S2** [failed claim] Can I just point out that every time we've seen this, it's been after a config change—
82.6–86.4 **S3** [turn] —no, last month it happened and nobody touched the configs! Remember? ▷ yields to S1
84.8–86.1 **S5** [laughter] Haha, exactly.
86.6–94.7 **S1** [turn] Look, we're going to have to check both logs sets or we'll just keep arguing. Can someone pull up the DNS and backend entries for this morning? ▷ yields to S4
94.6–99.3 **S4** [turn] I'll do that. But if we keep second-guessing every step, we're never getting out of this loop. ▷ yields to S5
95.6–96.4 **S2** [backchannel] Agreed.
99.3–101.8 **S5** [turn] Yeah, well, better than another outage next week. ▷ yields, end

dev split · tone: curious and inquisitive · noise: office hvac, 31.8 dB SNR; dishes clinking · voices: S1 M, S2 M, S3 M, S4 F, S5 F
EOT 17 yields / 0 holds · INT 1 interruptions / 4 negatives (4 backchannels, 0 laughter) · 4 failed claims (excluded)

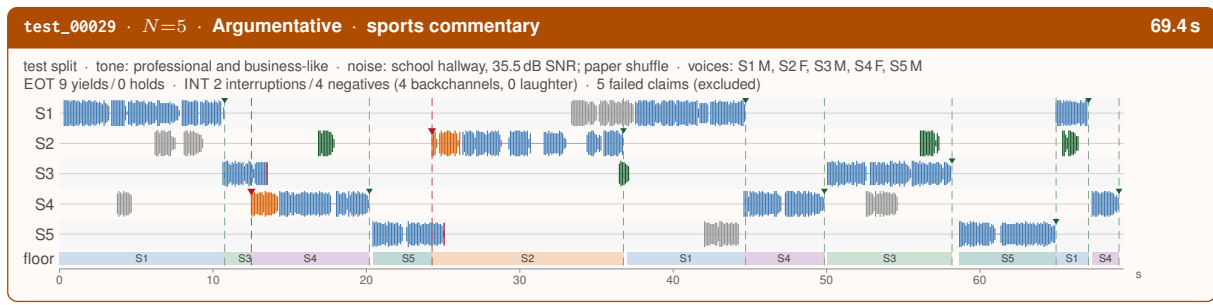


Transcript of dev_00032

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.1 **S5** [turn] All right, everyone. Today we're debating whether standardized testing is a fair way to measure student ability. Who wants to start us off? ▷ yields to S1
8.5–16.6 **S1** [turn] Yeah, I'll go first. I honestly think standardized tests are a terrible measure. They just— they only test memorization, not actual understanding.
12.3–12.8 **S3** [backchannel] mm-hm

13.9–15.6 **S2** [failed claim] But that's not strictly true—
 16.7–22.9 **S1** [turn] —let me finish, man. Like, literally, you could cram the night before and get a good score. That doesn't mean you actually get it. ▷ yields to S2
 22.9–32.6 **S2** [turn] What I was trying to say is—sure, they focus on knowledge recall, but there's clear data correlating high test scores with later academic success. ▷ yields to S4
 32.6–38.6 **S4** [turn] Okay, but correlation doesn't mean causation. Like, maybe students with more resources just test better because they can afford tutors and prep courses. ▷ cut off
 38.2–42.0 **S2** [interruption] That's a point, but if you control for socioeconomic status— ▷ takes floor from S4
 38.7–39.2 **S3** [backchannel] yeah
 42.1–47.3 **S2** [turn] —the predictive validity still holds, at least according to the College Board studies. ▷ yields to S1
 47.3–51.5 **S1** [turn] I don't buy that. Those studies are paid for by the same people selling the tests. ▷ yields to S5
 51.8–60.7 **S5** [turn] Let's not get conspiratorial—let's refocus. What about the argument that standardized tests encourage teaching to the test, limiting creativity? ▷ yields to S4
 60.6–67.5 **S4** [turn] See, that's exactly it. My high school experience was just constant drilling. I didn't feel like I learned how to think, just how to fill bubbles. ▷ yields to S2
 63.2–63.9 **S3** [backchannel] oh yeah
 67.5–72.7 **S2** [turn] But wait, is that a problem with the test, or is it about the way schools use them? ▷ yields to S1
 69.7–74.6 **S1** [failed claim] Dude, the test dictates how schools teach, you can't separate them. If the stakes are high—
 72.8–73.8 **S2** [failed claim] Not always—
 74.7–77.5 **S1** [turn] —they're gonna focus on that. It becomes a game. ▷ yields to S3
 78.0–83.4 **S3** [turn] Does anyone actually like tests? Because I just fill in random circles after an hour. ▷ yields to S4
 83.4–88.6 **S4** [turn] Haha, I mean, exactly. It's so stressful, too—like people literally get sick before SAT day. ▷ yields to S5
 85.3–85.8 **S1** [backchannel] yeah
 88.9–93.2 **S5** [turn] So, what's a better alternative? Portfolio assessments? Project-based learning? ▷ yields to S3
 93.1–97.5 **S3** [turn] Portfolios are subjective, though. Like, way more open to bias.
 96.0–98.9 **S2** [failed claim] Right, at least standardized tests are standardized—
 97.6–100.2 **S3** [turn] —and projects might depend on your teacher's mood. ▷ yields to S5
 100.7–106.9 **S5** [turn] I see a lot of strong opinions. Maybe the answer isn't either-or, but a combination. Okay, last thoughts? ▷ yields to S4
 106.8–111.5 **S4** [turn] If you're measuring people, you have to be fair. I just don't think tests alone can ever do that. ▷ yields to S2
 111.5–117.7 **S2** [turn] I agree, but we need some kind of objective measure, otherwise admissions just gets even more random. ▷ yields to S1
 117.7–122.1 **S1** [turn] Well, whatever we do now, it's definitely not working. Something's gotta change. ▷ yields, end

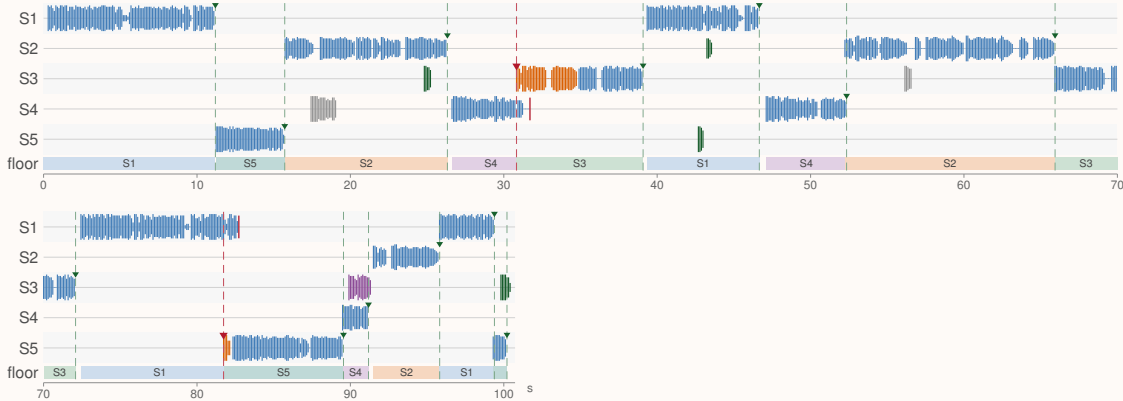


Transcript of test_00029

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–7.8 **S1** [turn] Alright, so everyone's talking about last night's game. I'm just gonna say it: the ref's call in the fourth quarter was absolutely ridiculous.
 3.8–4.7 **S4** [failed claim] Wait, hold on—
 6.2–9.3 **S2** [failed claim] Sure, blame the ref, classic analysis.
 8.0–10.8 **S1** [turn] —I mean, did you see that? There was barely even contact! ▷ yields to S3
 10.7–13.5 **S3** [turn] I, um, I think the contact was there, just maybe not enough for a flagrant— ▷ cut off
 12.5–14.3 **S4** [interruption] No, it was a clear foul, come on. ▷ takes floor from S3
 14.4–20.2 **S4** [turn] If you look at the replay, you see his elbow hit the shooter right on the arm. By the rule book, that's a foul every time. ▷ yields to S5
 16.9–17.9 **S2** [backchannel] Yeah, exactly.
 20.4–25.1 **S5** [turn] I, uh, but don't you think context matters? Like, late in the game, you, you maybe let them play a little— ▷ cut off
 24.3–26.2 **S2** [interruption] No, that's a slippery slope. ▷ takes floor from S5
 26.3–35.3 **S2** [turn] If you start officiating by situation instead of by the rulebook, consistency's gone. It's chaos.
 33.4–37.5 **S1** [failed claim] But they do it all the time, come on. Late-game, playoff, the whistle gets swallowed—
 35.5–36.8 **S2** [turn] —Doesn't make it right. ▷ yields to S1
 36.5–37.2 **S3** [backchannel] Right.
 37.6–44.7 **S1** [turn] Okay, so then let me ask you: if that's a foul, why do we see it get ignored half the season? Like, I've seen way worse just get waved off. ▷ yields to S4
 42.0–44.4 **S5** [failed claim] Consistency's a problem, I agree with that.
 44.6–49.9 **S4** [turn] But that's not the point, the point is what the rule says. We can't have every ref inventing their own standards. ▷ yields to S3
 50.1–55.5 **S3** [turn] I'm, uh, not sure the rule is as clear-cut as you think, I mean, isn't there, like, room for interpretation?
 52.6–54.7 **S4** [failed claim] No, it's written right there in the manual—
 55.6–58.2 **S3** [turn] —Yeah, but practice is different than theory, isn't it? ▷ yields to S5
 56.1–57.4 **S2** [backchannel] Welcome to sports.
 58.7–65.0 **S5** [turn] Can I just say, uh, I think both sides have a point. Maybe what we need is, well, better training for refs, not just tighter rules? ▷ yields to S1
 65.0–67.1 **S1** [turn] Sure, but good luck getting the league to pay for that. ▷ yields to S4
 65.4–66.5 **S2** [backchannel] Yeah, dream on.
 67.3–69.1 **S4** [turn] We can agree it's a mess, at least. ▷ yields, end

test split · tone: professional and business-like · noise: school hallway, 34.2 dB SNR; paper shuffle; distant speech · voices: S1 F, S2 F, S3 F, S4 M, S5 M
EOT 13 yields / 0 holds · INT 2 interruptions / 5 negatives (4 backchannels, 1 laughter) · 2 failed claims (excluded)

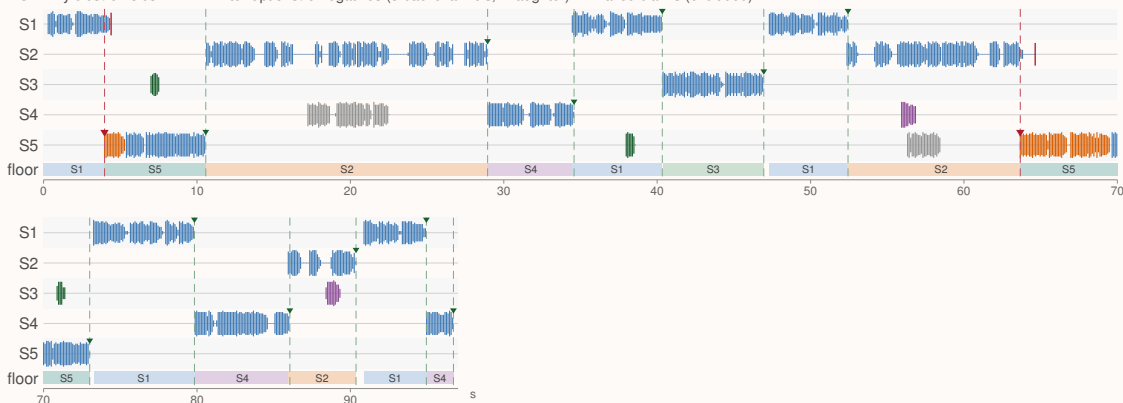


Transcript of test_00039

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–11.2 **S1** [turn] Alright, can we start by agreeing that someone needs to pick up Mom after her appointment at three? I just want to make sure we're all on the same page because last week was, uh, kind of chaotic. ▷ yields to S5
11.2–15.7 **S5** [turn] Well, honestly, I can do it if no one minds me shifting dinner a bit later again. ▷ yields to S2
15.7–20.3 **S2** [turn] But that pushes back our whole evening. I really don't think it's sustainable every week—
17.4–19.1 **S4** [failed claim] Hang on, can I just interject—
20.4–26.3 **S2** [turn] —to have dinner at eight thirty every Thursday just because the timing isn't convenient for anyone else. ▷ yields to S4
24.8–25.3 **S3** [backchannel] Right
26.6–31.7 **S4** [turn] If I leave work a touch early, I could squeeze it in, but traffic's honestly brutal at three. Like if somebody else could cover today, I'd take next week? ▷ cut off
30.8–34.8 **S3** [interruption] Wait, but what if we just call her a rideshare? She's done it before, and I mean— ▷ takes floor from S4
34.9–39.1 **S3** [turn] —she's totally fine with it, she said she actually liked chatting with her last driver! ▷ yields to S1
39.3–46.7 **S1** [turn] But she called me after and said it stressed her out not knowing who would show up, so I'd really prefer if it's one of us. I get her point. ▷ yields to S4
42.7–43.1 **S5** [backchannel] Yeah
43.2–43.6 **S2** [backchannel] Mm-hm
47.1–52.3 **S4** [turn] But if she's that stressed, why not move her appointment to a better time? Has anyone even asked her about that? ▷ yields to S2
52.2–60.0 **S2** [turn] She can't, the doctor's booked months out and she doesn't want to reschedule again. Look, this shouldn't fall on just two people every time—
56.1–56.6 **S3** [failed claim] But—
60.1–65.9 **S2** [turn] —last time I had to rearrange my entire afternoon. Sorry, but we've got to be more fair here. ▷ yields to S3
65.9–72.1 **S3** [turn] Okay, but I literally can't get out of rehearsal before four—it's not me bailing, it's just not possible. ▷ yields to S1
72.4–82.7 **S1** [turn] Alright, so it sounds like today it's down to either me or, um, Greg, and I can move my meeting if it means she feels comfortable. But next week let's formalize a rotation, so it's not always up for debate— ▷ cut off
81.7–82.2 **S5** [interruption] Wait— ▷ takes floor from S1
82.3–89.5 **S5** [turn] If you're already moving stuff, let me just take this one, honestly, it's not a huge deal for me. Next week, rotation though, for sure. ▷ yields to S4
89.5–91.2 **S4** [turn] Thank you, Greg. That helps a lot. ▷ yields to S2
89.9–91.4 **S3** [laughter] haha, crisis averted!
91.5–95.8 **S2** [turn] Yeah, thanks. Can we actually put the rotation schedule in writing this time? ▷ yields to S1
95.8–99.4 **S1** [turn] Absolutely, I'll send an email out tonight. And, thank you, Greg. ▷ yields to S5
99.3–100.2 **S5** [turn] No problem at all. ▷ yields, end
99.8–100.4 **S3** [backchannel] Perfect

test split · tone: enthusiastic and energetic · noise: outdoor park, 37.2 dB SNR; chairs scraping · voices: S1 M, S2 F, S3 F, S4 F, S5 F
EOT 12 yields / 0 holds · INT 2 interruptions / 5 negatives (3 backchannels, 2 laughter) · 2 failed claims (excluded)



Transcript of test_00079

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–4.4 **S1** [turn] All right, if we're being honest, last night's game was a textbook example of why you don't trust analytics over your gut, especially in the final two

minutes. > cut off

4.0–5.3 **S5** [interruption] Oh come on, that's such— > takes floor from S1

5.4–10.6 **S5** [turn] —a tired argument. If they'd listened to gut, they'd have benched their top scorer again and lost by even more. > yields to S2

7.0–7.6 **S3** [backchannel] mm-hm

10.6–20.4 **S2** [turn] Yeah, and let's not pretend the data was telling them to put Carter on defense. That was pure panic mode.

17.2–22.6 **S4** [failed claim] But did you see Carter's face? He looked like he'd never seen a basketball before, I mean, honestly—

20.5–28.9 **S2** [turn] —yeah, that's on coaching, not him! You throw someone in, you give them support, not just a stat sheet. > yields to S4

29.0–34.6 **S4** [turn] Still, if you're a pro, you gotta handle pressure. He fumbled two passes! That's elementary stuff. > yields to S1

34.5–40.3 **S1** [turn] Exactly! There's no metric for freezing up. That's human nature, that's why you need guys who've been in the trenches. > yields to S3

38.0–38.6 **S5** [backchannel] Come on—

40.4–46.9 **S3** [turn] But isn't it also true the veterans had way more turnovers this season? Benching them was justifiable, wasn't it? > yields to S1

47.3–52.4 **S1** [turn] Turnovers, yeah, but you trust someone with playoff nerves, not some rookie who's just happy to be there. > yields to S2

52.4–58.5 **S2** [turn] That's funny, since last year you said young legs win games in crunch time.

55.9–56.9 **S4** [laughter] Oh wow, receipts!

56.3–58.5 **S5** [failed claim] Let's get back to strategy for a second—

58.6–64.6 **S2** [turn] —because you can't say one thing in May and the total opposite in June. That's all I'm saying. > cut off

63.7–69.5 **S5** [interruption] Okay, but if you're down by four, why not foul and get the ball back? Analytics or not, it's basic clock management— > takes floor from S2

69.7–73.0 **S5** [turn] —instead they let the clock bleed, gave up a wide-open three, game over. > yields to S1

70.9–71.4 **S3** [backchannel] Yeah...

73.3–79.8 **S1** [turn] That's on the coaching staff, I'll give you that one. But you can't reduce everything to a spreadsheet. At some point, heart matters, too. > yields to S4

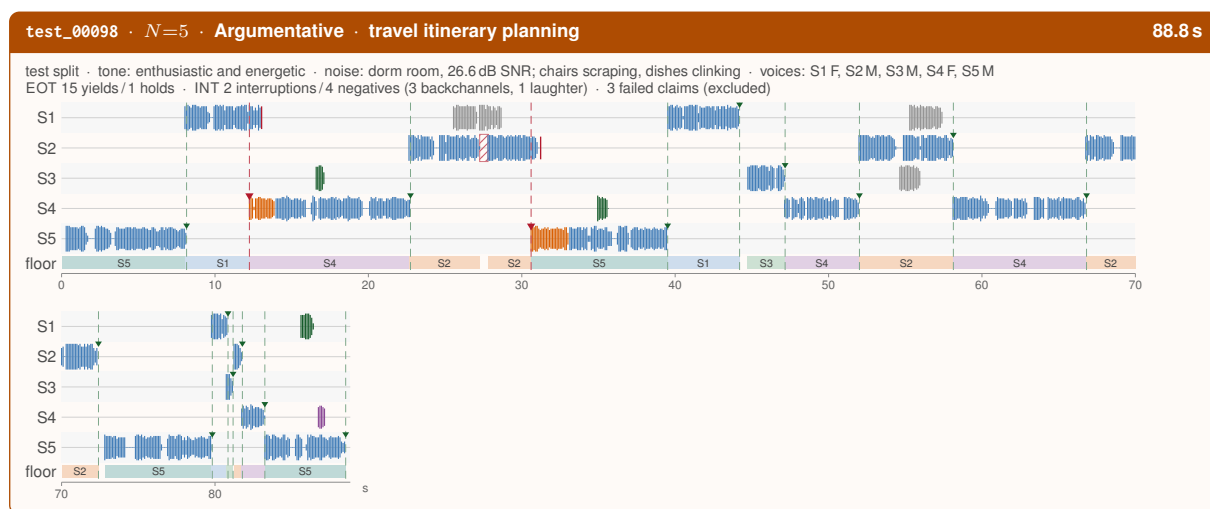
79.9–86.1 **S4** [turn] But if you're gonna talk heart, you can't ignore how badly Simmons wanted that shot at the end. He was calling for it! > yields to S2

86.0–90.4 **S2** [turn] And he bricked it, but sure, let's call that leadership. > yields to S1

88.4–89.3 **S3** [laughter] haha right

90.9–95.0 **S1** [turn] I still say you put players in position to win, not to run out the clock on hope. > yields to S4

95.0–96.7 **S4** [turn] But it was fun to watch, so there's that! > yields, end



Transcript of test_00098

start–end (s) speaker [ground-truth label] text > outcome

0.3–8.2 **S5** [turn] Alright, let's get started! We've only got a week, so we need to settle on where we're actually going and what we're prioritizing, okay? > yields to S1

8.0–13.0 **S1** [turn] Paris. It has to be Paris! There's just so much to do and see, like, I have a whole list of stuff— > cut off

12.2–13.9 **S4** [interruption] Wait, but you always say Paris! > takes floor from S1

14.0–22.7 **S4** [turn] Can we do something actually different for once? Like, can we please consider Lisbon or, I don't know, Prague? Something not Paris, just for the love of croissants— > yields to S2

16.6–17.2 **S3** [backchannel] uh, yeah

22.6–27.3 **S2** [turn] I mean, Lisbon is interesting but I think flights will be pricier this close to summer. > holds, pauses

25.6–28.7 **S1** [failed claim] But Paris is just classic, right? Also, public transport—

27.8–31.2 **S2** [turn] —And, honestly, Prague has cheaper everything and it's gorgeous in June. It's— > cut off

30.6–33.1 **S5** [interruption] Hold on, can we stick to one city at a time? > takes floor from S2

33.1–39.5 **S5** [turn] Let's do Paris first, then argue about the others. So, uh, who's actually in favor of Paris, hands up— > yields to S1

35.0–35.7 **S4** [backchannel] Oh come on!

39.5–44.2 **S1** [turn] Me, absolutely, and you all are gonna thank me when we're eating cr textasciicircumpeys by the Seine at midnight. > yields to S3

44.7–47.2 **S3** [turn] I, um, I actually would rather do Prague. Sorry. > yields to S4

47.1–52.0 **S4** [turn] See? Thank you! Prague! The jazz bars and that weird clock tower? Way more interesting! > yields to S2

52.0–58.1 **S2** [turn] Prague's affordable, but what about the language barrier? Isn't that a problem for, uh, you know, ordering food, trains— > yields to S4

54.6–56.0 **S3** [failed claim] Yeah, but not everywhere, I think...

55.3–57.4 **S1** [failed claim] Everyone speaks English nowadays, come on.

58.1–66.8 **S4** [turn] It's way more of an adventure not knowing everything! Paris is, like, tourist central. No offense, but do you really want crowds and lines for everything? > yields to S2

66.8–72.4 **S2** [turn] But the food in Paris is unbeatable. Also, the museums—if we plan well, we can skip some lines... > yields to S5

72.8–79.8 **S5** [turn] Okay, time out, time out. We're all just circling, so let's vote: Paris, Prague, or Lisbon, one word only, go. > yields to S1

79.8–80.8 **S1** [turn] Paris, obviously. > yields to S3

80.7–81.2 **S3** [turn] Prague. > yields to S2

81.2–81.8 **S2** [turn] Prague. > yields to S4

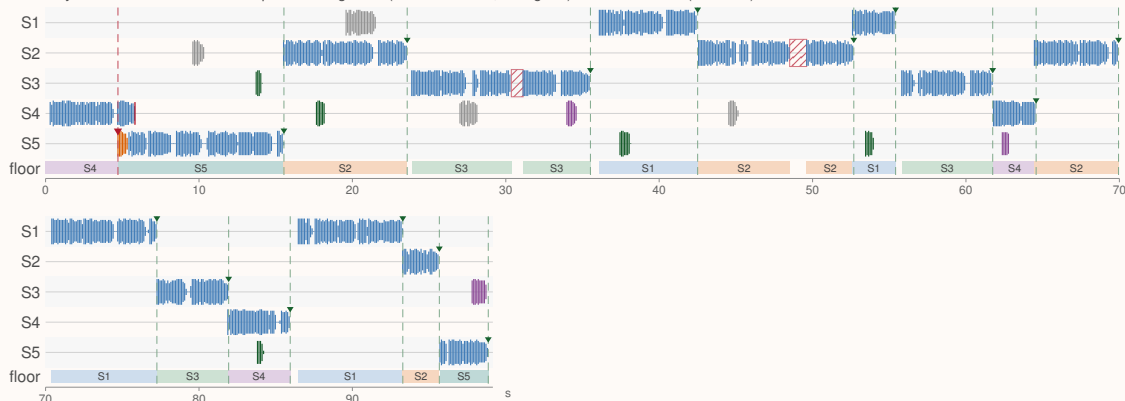
81.7–83.2 **S4** [turn] Lisbon! Wait, seriously?! > yields to S5

83.3–88.5 **S5** [turn] Looks like Prague wins, folks. Next: hotels or hostels—brace yourselves. > yields, end

85.6–86.4 **S1** [backchannel] Nooo, come on!

86.7–87.2 **S4** [laughter] haha

test split · tone: casual and friendly · noise: office hvac, 22.7 dB SNR; alarm beep, door knock · voices: S1 F, S2 M, S3 M, S4 F, S5 F
EOT 15 yields / 2 holds · INT 1 interruptions / 8 negatives (5 backchannels, 3 laughter) · 4 failed claims (excluded)



Transcript of test_00142

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–5.8 **S4** [turn] Okay, hear me out, I honestly think homework should not even count for grades, like, people have lives and jobs and— ▷ cut off

4.7–5.4 **S5** [interruption] But wait! ▷ takes floor from S4

5.4–12.5 **S5** [turn] If homework didn't count, I would literally never do it. I know myself, I'd just, like... binge Netflix nonstop instead.

9.6–10.3 **S2** [failed claim] Yeah but—

12.6–15.5 **S5** [turn] —so it has to matter, or no one's gonna bother. Right? ▷ yields to S2

13.7–14.1 **S3** [backchannel] mm-hm

15.5–23.6 **S2** [turn] But see, um, if homework is too much of your grade, then it punishes people who have, y'know, other responsibilities. Like, what if someone works every night? ▷ yields to S3

17.7–18.3 **S4** [backchannel] Totally!

19.6–21.5 **S1** [failed claim] Exactly, that's such a good point.

23.9–30.4 **S3** [turn] I mean, but isn't the real problem how much busywork some teachers assign? Like, if the homework felt meaningful instead— ▷ holds, pauses

27.0–28.2 **S4** [failed claim] See! That's what I'm saying!

31.1–35.5 **S3** [turn] —then maybe people wouldn't resent it so much. It's not the grade, it's the pointlessness. ▷ yields to S1

34.0–34.6 **S4** [laughter] haha right

36.1–42.5 **S1** [turn] I get that, but for some kids, like, the only reason they're passing is the homework points. If you take those away, what happens to them? ▷ yields to S2

37.4–38.2 **S5** [backchannel] Whoa, true.

42.6–48.5 **S2** [turn] But then does that mean the system's just set up for, uh, like, checking boxes instead of actually learning anything? ▷ holds, pauses

44.5–45.1 **S4** [failed claim] No, but—

49.6–52.7 **S2** [turn] —like, should getting the grade just mean you followed instructions? ▷ yields to S1

52.6–55.4 **S1** [turn] But sometimes following instructions is the point, isn't it? ▷ yields to S3

53.5–54.0 **S5** [backchannel] oh wow

55.8–61.7 **S3** [turn] I think it's both. Like, grades should show effort and understanding, not just one. But that's hard to measure, right? ▷ yields to S4

61.8–64.6 **S4** [turn] Yeah but that just makes it all so subjective, doesn't it? ▷ yields to S2

62.4–62.9 **S5** [laughter] haha

64.5–69.9 **S2** [turn] Well, I mean, some teachers are super fair and others just, like, give points for anything or nothing, so... yeah. ▷ yields to S1

70.4–77.3 **S1** [turn] That's why I think, um, having some homework count but not too much is a good middle ground. Maybe ten percent or something, not like... fifty. ▷ yields to S3

77.3–81.9 **S3** [turn] Right, a little keeps people motivated, but you don't get sunk if you miss a couple assignments. ▷ yields to S4

81.9–85.9 **S4** [turn] Okay, but what if you do all the homework and still bomb the tests though? Then what? ▷ yields to S1

83.8–84.2 **S5** [backchannel] mm-hm

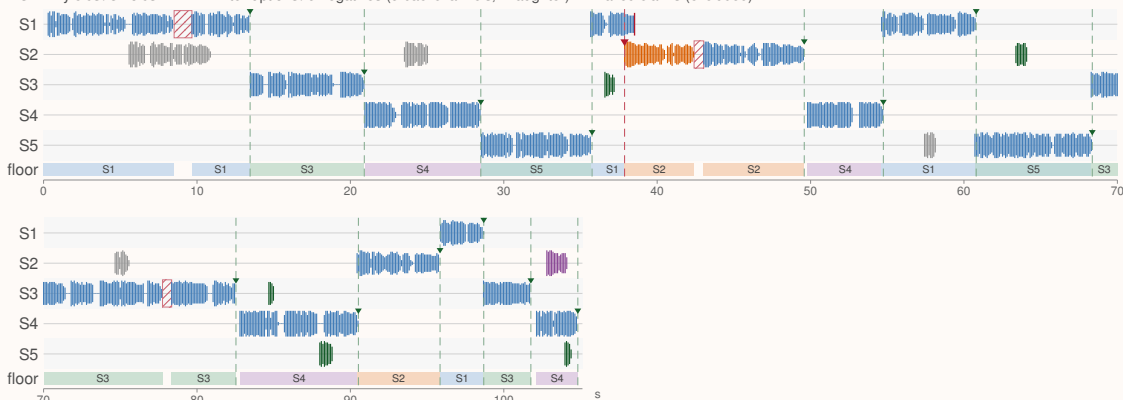
86.5–93.3 **S1** [turn] Yeah, that's rough. Maybe there should be like, a retake policy or something, so homework preps you but doesn't completely determine your grade. ▷ yields to S2

93.3–95.7 **S2** [turn] I wish every class did that, honestly. ▷ yields to S5

95.7–98.9 **S5** [turn] Same! I would've had way less panic attacks in Algebra two. ▷ yields, end

97.8–98.8 **S3** [laughter] haha same here

test split · tone: professional and business-like · noise: school hallway, 29.7 dB SNR; paper shuffle · voices: S1 F, S2 M, S3 F, S4 M, S5 M
EOT 14 yields/3 holds · INT 1 interruptions/6 negatives (5 backchannels, 1 laughter) · 4 failed claims (excluded)



Transcript of test_00147

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–8.5 **S1** [turn] Let's stop dancing around it—last night's decision to bench Carter in the fourth was indefensible. You don't sit your star when you're down two in a playoff game. ▷ holds, pauses

5.6–10.9 **S2** [failed claim] But—look, that's just it! He was ice cold by then, Steph, he had what, four turnovers in the third alone?

9.7–13.5 **S1** [turn] —You play through it. He makes things happen under pressure, even off nights. ▷ yields to S3

13.5–20.9 **S3** [turn] Can I just add, both points are valid, but benching Carter signaled panic from Coach Ellis. That rattled the rest of the lineup. ▷ yields to S4

20.9–26.4 **S4** [turn] Right, and I'd argue the whole vibe changed. You could see the offense just stall, even their body language—

23.5–25.1 **S2** [failed claim] Sure, the offense stalled, but—

26.6–28.5 **S4** [turn] —was off. The confidence just sapped out. ▷ yields to S5

28.5–35.8 **S5** [turn] Um, if—if I can weigh in, I think, uh, benching him might've been about defense more than offense? Maybe Coach Ellis was— ▷ yields to S1

35.7–38.5 **S1** [turn] That doesn't hold up. Carter's not a liability on D, not in crunch time. ▷ cut off

36.6–37.3 **S3** [backchannel] Exactly.

37.9–42.4 **S2** [interruption] But—hang on, look, Lewis was burning him every switch. Just check the film, it's right there. ▷ takes floor from S1 ▷ holds, pauses

43.0–49.6 **S2** [turn] I'm not saying Carter isn't your closer, but on a night like that? You adjust. It's not about star power, it's about matchups in those moments. ▷ yields to S4

49.8–54.7 **S4** [turn] But then who do you trust more than Carter to create out of nothing? Nobody else wants the ball there. ▷ yields to S1

54.6–59.5 **S1** [turn] Exactly my point. You live and die by your best. Anything else invites second-guessing.

57.4–58.2 **S5** [failed claim] I—um, sorry—

59.6–60.8 **S1** [turn] —Go ahead, Max. ▷ yields to S5

60.7–68.4 **S5** [turn] Oh, uh, thanks. I, I just think there's blame on Ellis for not calling time sooner too. Team looked totally lost when the run started. ▷ yields to S3

63.4–64.1 **S2** [backchannel] Right, right.

68.3–77.8 **S3** [turn] But also, they've leaned so hard on Carter all season, you had to expect fatigue, and, honestly, that's on the front office for depth issues, not just coaching. ▷ holds, pauses

74.7–75.6 **S2** [failed claim] I don't buy that—

78.3–82.5 **S3** [turn] —You can't run a seven-man rotation in May and not expect breakdowns. ▷ yields to S4

82.8–90.5 **S4** [turn] Can we agree though, the officiating in the fourth didn't help? Carter drew contact twice, no whistle. That stuff impacts decision-making too. ▷ yields to S2

84.7–85.1 **S3** [backchannel] Mm-hm.

88.0–88.9 **S5** [backchannel] Yeah, definitely.

90.4–95.8 **S2** [turn] Okay, but even with bad calls, if you put Carter back with four left, maybe we're not even having this debate. ▷ yields to S1

95.9–98.7 **S1** [turn] Exactly. Ellis overthought it, plain and simple. ▷ yields to S3

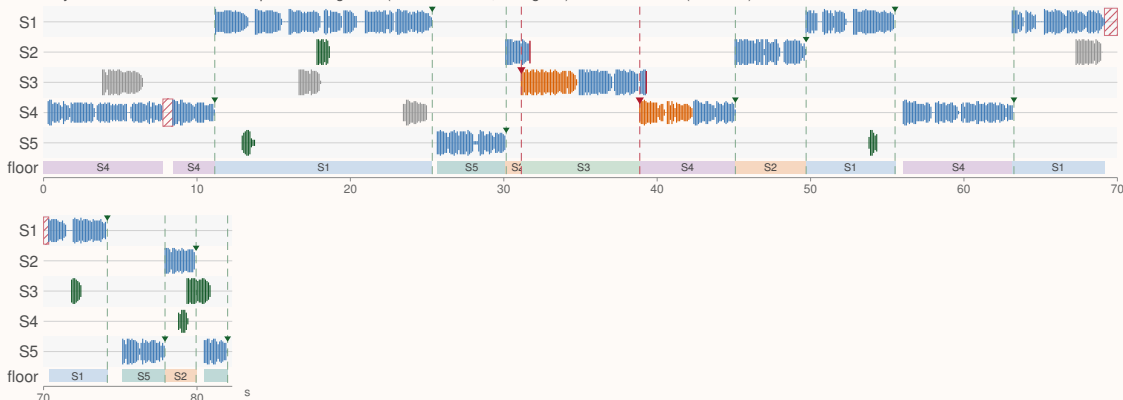
98.7–101.8 **S3** [turn] I still say it's a symptom of a bigger problem with their bench. ▷ yields to S4

102.1–104.8 **S4** [turn] I'd bet there's a shakeup this summer if they go out early. ▷ yields, end

102.8–104.1 **S2** [laughter] Haha, you might be right.

104.0–104.4 **S5** [backchannel] Yeah.

test split · tone: professional and business-like · noise: cafe, 23.4 dB SNR; phone ring, keyboard clatter · voices: S1 F, S2 M, S3 M, S4 M, S5 M
EOT 11 yields/2 holds · INT 2 interruptions/6 negatives (6 backchannels, 0 laughter) · 4 failed claims (excluded)



Transcript of test_00181

start—end (s) speaker [ground-truth label] text > outcome

0.3–7.8 **S4** [turn] Okay, team, next question: What year did the Berlin Wall fall? I'm pretty sure it's 1989, right? I mean, I remember my dad talking about it— > holds, pauses

3.8–6.5 **S3** [failed claim] Wait, uh, wasn't it ninety or maybe ninety-one?

8.4–11.2 **S4** [turn] —like, it was definitely eighties, not nineties, I think. > yields to S1

11.2–19.5 **S1** [turn] Actually, it is 1989. The wall came down that November. The official reunification was ninety, but the fall itself—

13.0–13.8 **S5** [backchannel] She's right.

16.7–18.1 **S3** [failed claim] But didn't the official—

17.8–18.7 **S2** [backchannel] Yeah, that's right.

19.6–25.3 **S1** [turn] —was eighty-nine. There's a distinction between when it fell and when Germany became one country again. > yields to S5

23.5–25.0 **S4** [failed claim] Exactly! That's what I meant.

25.7–30.2 **S5** [turn] Alright, next: Which planet is closest to the sun? It's Mercury, right? Am I wrong? > yields to S2

30.1–31.7 **S2** [turn] Yeah, Mercury is the closest. > cut off

31.1–34.8 **S3** [interruption] Uh, well, I heard Venus sometimes gets closer because of orbital— > takes floor from S2

34.9–39.3 **S3** [turn] Like, there's this thing with elliptical orbits, I don't remember the details, but maybe Venus gets closer for part of its orbit? > cut off

38.9–42.3 **S4** [interruption] No way, it's always Mercury. That's like basic astronomy. > takes floor from S3

42.4–45.1 **S4** [turn] Seriously, it's Mercury. Venus is after that. > yields to S2

45.1–49.7 **S2** [turn] Right, Mercury never leaves that spot closest to the sun, even with the weird orbits. > yields to S1

49.7–55.5 **S1** [turn] To clarify, Venus gets closer to Earth, but Mercury is always the closest to the sun itself. > yields to S4

53.8–54.4 **S5** [backchannel] Mhm.

56.0–63.3 **S4** [turn] Alright, let's move on before we spiral. What about the Shakespeare question? Was it Hamlet or Macbeth that has the line about 'sound and fury'? > yields to S1

63.1–69.2 **S1** [turn] That is Macbeth, act five, 'a tale told by an idiot, full of sound and fury, signifying nothing.' > holds, pauses

67.3–69.0 **S2** [failed claim] I thought Hamlet had some similar—

70.3–74.2 **S1** [turn] —not that line, no. Hamlet's existential stuff is different. > yields to S5

71.8–72.5 **S3** [backchannel] Yeah.

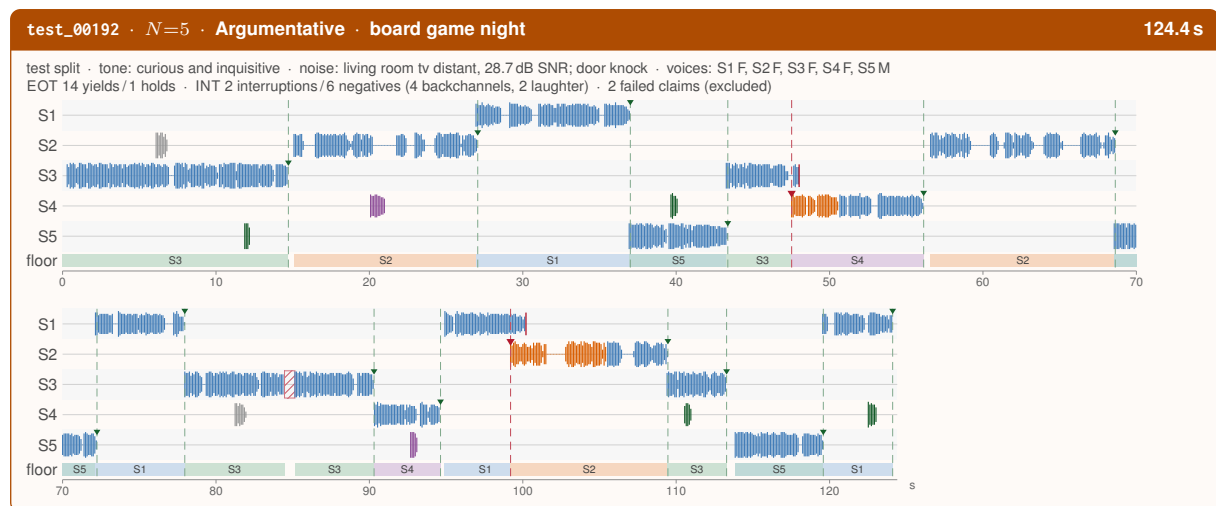
75.2–77.9 **S5** [turn] Quick, bonus round: which country invented paper? > yields to S2

77.9–79.9 **S2** [turn] China. Around two thousand years ago. > yields to S5

78.8–79.4 **S4** [backchannel] Agreed.

79.4–80.9 **S3** [backchannel] Yeah, definitely China.

80.5–82.0 **S5** [turn] Alright, let's lock those in. > yields, end



Transcript of test_00192

start—end (s) speaker [ground-truth label] text > outcome

0.3–10.1 **S3** [turn] Okay look, Settlers of Catan is objectively the best game for a group because everyone's actually involved every turn, even if it's just trading or blocking somebody.

6.1–6.8 **S2** [failed claim] Yeah but—

10.2–14.7 **S3** [turn] —and you can't just zone out for half the game like you do with Ticket to Ride! I'm just saying. > yields to S2

11.9–12.3 **S5** [backchannel] mm-hm

15.1–27.1 **S2** [turn] No, see, Catan is just a resource math spreadsheet with sheep. It's, like, Monopoly for people who pretend they don't like Monopoly. > yields to S1

20.1–21.1 **S4** [laughter] Haha, oh my god!

27.0–37.0 **S1** [turn] But that's actually the point, right? The trading makes it social. With Ticket to Ride, you just draw cards and hope no one cuts your track. There's barely any interaction. > yields to S5

36.9–43.4 **S5** [turn] I mean, have you ever played Risk with this group? That's way more interactive, like, there's open betrayal almost instantly. > yields to S3

39.7–40.1 **S4** [backchannel] Right

43.3–48.0 **S3** [turn] Yes, but Risk takes six hours and ends with us not speaking for a week. We're still not allowed to talk about Australia, remember? > cut off

47.5–50.6 **S4** [interruption] Wait, but what about, um, Betrayal at House on the Hill? > takes floor from S3

50.7–56.1 **S4** [turn] I mean, it's totally unpredictable and everyone gets their little moment to be a chaos gremlin. > yields to S2

56.6–68.8 **S2** [turn] But the rules literally change halfway and then nobody knows what's happening except maybe Ben, and he gets so smug when he does. > yields to S5

68.6–72.2 **S5** [turn] Hey, it's not my fault the haunt is always confusing. That's the charm. > yields to S1

72.1–78.0 **S1** [turn] But if clarity matters, honestly Splendor is just pure strategy and no drama. It's elegant. > yields to S3

78.0–84.5 **S3** [turn] Elegant but boring. I've never seen anyone actually argue during Splendor, which is... maybe a bad sign for tonight? > holds, pauses

81.3–82.0 **S4** [failed claim] Oh come on—

85.1–90.3 **S3** [turn] —like, you just collect gems and occasionally glare at whoever took your card. That's not the vibe. > yields to S4

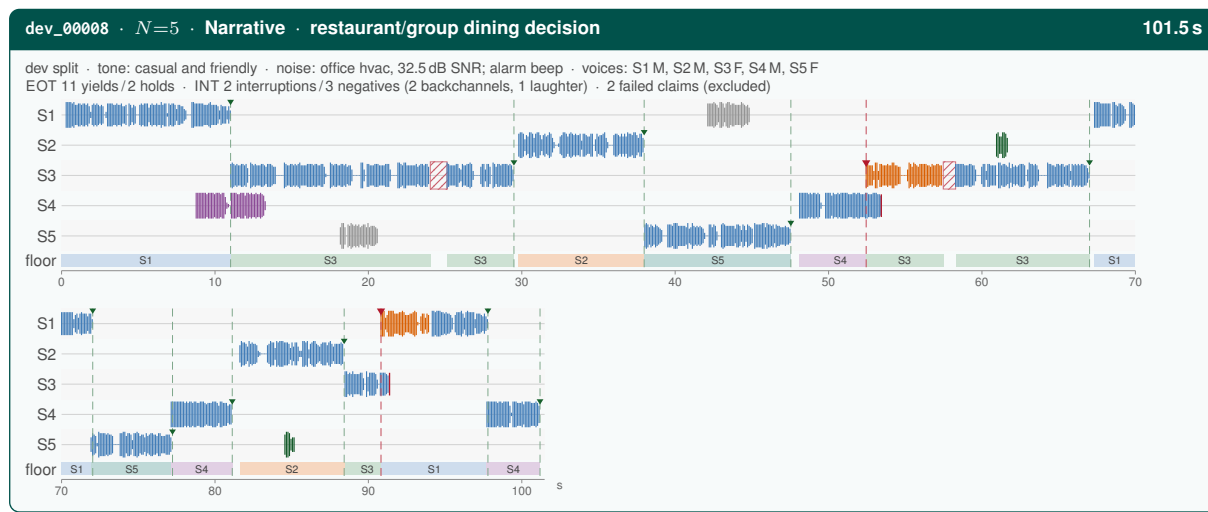
90.3–94.6 **S4** [turn] But sometimes you need a chill game or you'll flip the table, Jessica, seriously. > yields to S1

92.7–93.2 **S5** [laughter] haha

94.9–100.2 **S1** [turn] Honestly, I think the ideal is something with strategy but still room for side-eye and trash talk. Codenames maybe? > cut off

99.2–105.5 **S2** [interruption] But nobody agrees about the clues! It's ten minutes of arguing every round— > takes floor from S1

105.5–109.5 **S2** [turn] —which, okay, maybe that’s actually what you all want. ▷ yields to S3
 109.4–113.3 **S3** [turn] Yeah, it’s more fun to bicker than quietly stack cards! You know it’s true. ▷ yields to S5
 110.6–111.0 **S4** [backchannel] Sure
 113.9–119.6 **S5** [turn] I can live with Codenames if we promise not to yell at the spymaster this time. No guarantees, though. ▷ yields to S1
 119.6–124.1 **S1** [turn] Deal. But only if we rotate the spymaster, so nobody gets a complex. ▷ yields, end
 122.5–123.1 **S4** [backchannel] Agreed!

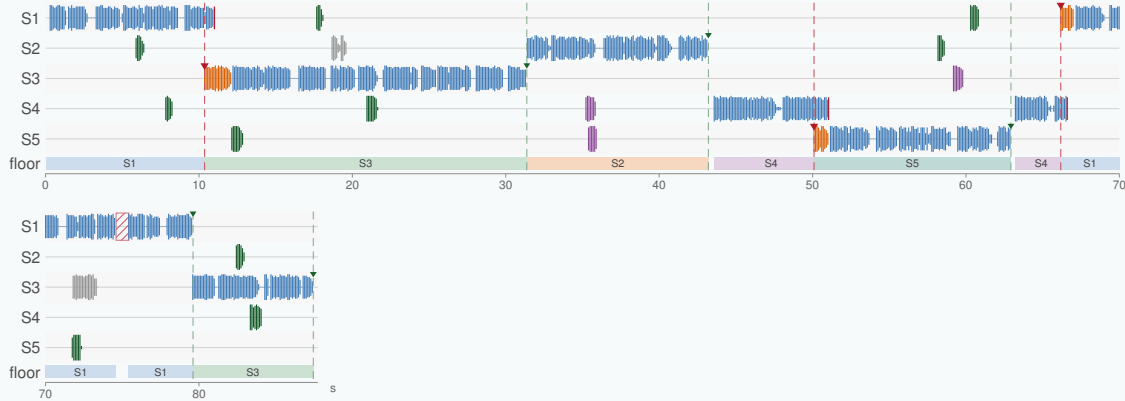


Transcript of dev_00008

start–end (s) speaker [ground-truth label] text ▷ outcome

0.3–11.0 **S1** [turn] All right, so before we do the whole ‘Where do you want to eat, I don’t know, where do you want to eat’ dance—can I just say, I refuse to go to that dim sum place again. Last time they literally forgot my order three times. ▷ yields to S3
 8.8–13.3 **S4** [laughter] Haha oh man, yeah, you looked so betrayed. That waitress would walk right past you every time.
 11.0–24.0 **S3** [turn] No, but seriously, that place is always chaos. I remember once they sat me and Sarah at a table for eight, and then tried to move us mid-meal. Like, with the soup already on the table. I had to balance the teapot in one hand— ▷ holds, pauses
 18.2–20.7 **S5** [failed claim] Right, and everyone just pretends that’s normal.
 25.1–29.5 **S3** [turn] —and they nearly dropped buns on my lap! So, uh, I can pass on that one too. ▷ yields to S2
 29.8–38.0 **S2** [turn] Noted. Avoid chaotic dim sum. But honestly, my only request is not pizza again. We do pizza literally every time. ▷ yields to S5
 38.0–45.1 **S5** [turn] Can I suggest ramen? The place over on Main does that spicy miso, and it’s, like, honestly the only thing I’ve craved all week.
 42.1–44.9 **S1** [failed claim] Didn’t you almost cry last time from the spice though?
 45.2–47.5 **S5** [turn] —Okay yes, but it’s a good pain. It’s worth it. ▷ yields to S4
 48.1–53.4 **S4** [turn] I mean, ramen does sound great. I do have this, uh, vivid memory though—last winter, I dropped an entire boiled egg in my lap right there. Still have the stain on my coat. ▷ cut off
 52.5–57.5 **S3** [interruption] Oh no, that place with the slippery benches? I think every table around us heard you curse. ▷ takes floor from S4 ▷ holds, pauses
 58.3–67.0 **S3** [turn] But their broth is amazing, so, honestly, accidental egg injuries aside, I’m on board. I can even call ahead and see if they’ll do a reservation. ▷ yields to S1
 61.0–61.7 **S2** [backchannel] Good call.
 67.3–72.0 **S1** [turn] Yeah, but there is a wait every Friday. Last time we waited what, like forty minutes? ▷ yields to S5
 71.9–77.2 **S5** [turn] True, but it’s so worth it. And last time we played that weird trivia game on your phone, remember? ▷ yields to S4
 77.1–81.1 **S4** [turn] Haha, you mean the one where Josh argued that octopuses can, like, read minds? ▷ yields to S2
 81.6–88.4 **S2** [turn] Still contend they might. But hey, as long as it isn’t pizza, I’ll wait for ramen, psychic octopus or not. ▷ yields to S3
 84.6–85.2 **S5** [backchannel] Yes!
 88.5–91.4 **S3** [turn] Great, so everyone’s good? I’ll call them now. Oh, wait— ▷ cut off
 90.8–94.0 **S1** [interruption] Wait, one vote for no mushrooms in the broth, please. ▷ takes floor from S3
 94.2–97.8 **S1** [turn] Every time they leave them in, I end up fishing them out like a raccoon. ▷ yields to S4
 97.7–101.2 **S4** [turn] Haha, noted, mushroom ban. We’ll make sure you get a clean bowl this time. ▷ yields, end

dev split · tone: casual and friendly · noise: outdoor park, 32.2 dB SNR; notification chime · voices: S1 F, S2 M, S3 M, S4 M, S5 F
 EOT 5 yields/1 holds · INT 3 interruptions/13 negatives (10 backchannels, 3 laughter) · 2 failed claims (excluded)



Transcript of dev_00017

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–11.0 **S1** [turn] Good evening everyone, and thanks for tuning in. You know, back in my coaching days, I actually saw a striker pull off a nutmeg like that in the state finals—exact same move we just watched—except the ball hit a sprinkler head and shot off in the wrong direction. That was a moment. ▷ cut off

5.9–6.4 **S2** [backchannel] Right.

7.8–8.3 **S4** [backchannel] yeah

10.4–12.1 **S3** [interruption] You know, if I can jump in— ▷ takes floor from S1

12.1–12.9 **S5** [backchannel] No way!

12.2–25.5 **S3** [turn] That reminds me of a match where the stats just didn't tell the full story. The striker had a low xG, but he kept creating space, pulling defenders out. Sometimes you only see that on the field, not in the numbers.

17.7–18.1 **S1** [backchannel] mm-hm

18.7–19.6 **S2** [failed claim] Yeah, but—

20.9–21.7 **S4** [backchannel] exactly

25.6–31.4 **S3** [turn] —it's those off-the-ball movements, right? They don't get a line on the stat sheet, but they win games. ▷ yields to S2

31.4–43.2 **S2** [turn] I was just gonna say, it, uh, reminds me of my old high school team. We had this one guy who barely touched the ball but just, like, ran everywhere. He looked exhausted after ten minutes. ▷ yields to S4

35.2–35.9 **S4** [laughter] haha right

35.4–36.0 **S5** [laughter] Haha

43.6–51.0 **S4** [turn] I used to play keeper, and I remember this one tryout where a kid took a shot from midfield. It floated so slowly, I had time to tie my shoe and still catch it. Sometimes it's just— ▷ cut off

50.1–51.0 **S5** [interruption] Can I just say— ▷ takes floor from S4

51.2–62.9 **S5** [turn] —I had almost the exact opposite moment last week! I was coaching the U10s and this little kid blasted a corner kick right into his own face. Everyone froze, but he just got up and laughed. Kids are so tough. ▷ yields to S4

58.2–58.7 **S2** [backchannel] mm-hm

59.2–59.9 **S3** [laughter] haha

60.3–60.9 **S1** [backchannel] oh wow

63.2–66.6 **S4** [turn] That's hilarious. You know, kids just bounce back. I broke a tooth at ten and played the rest of the game— ▷ cut off

66.2–67.1 **S1** [interruption] Wait, you what? ▷ takes floor from S4

67.2–74.6 **S1** [turn] —You finished the game with a broken tooth? That's some real dedication. I tell my youth players not to do that, but honestly, I get it. ▷ holds, pauses

71.7–72.3 **S5** [backchannel] so tough

71.8–73.4 **S3** [failed claim] You know, it's interesting—

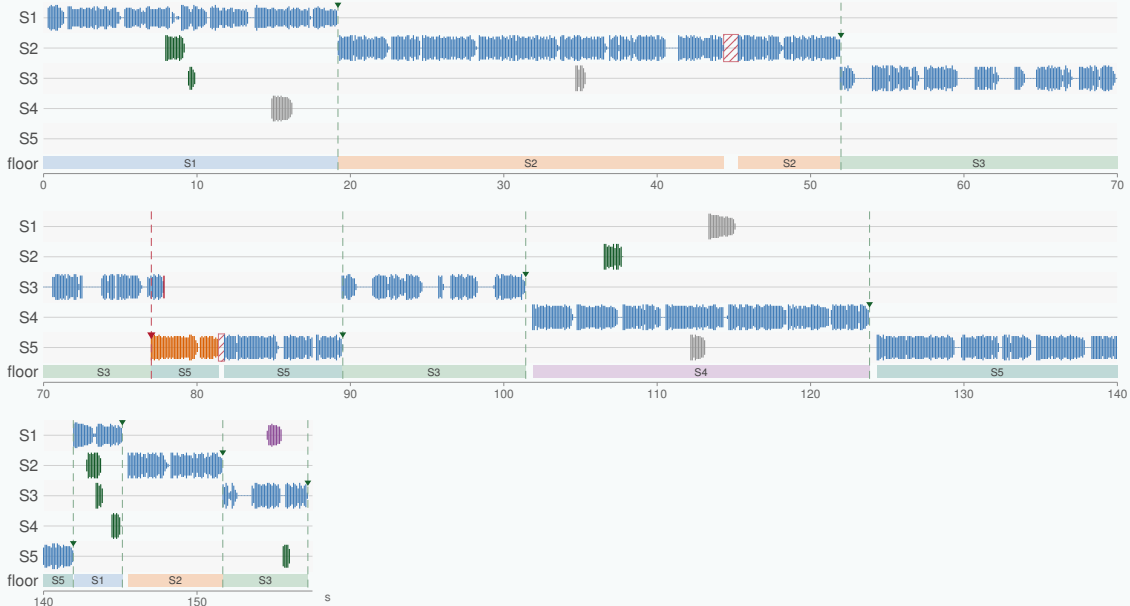
75.4–79.6 **S1** [turn] —resilience like that, it's hard to teach. Players either have it or they don't. ▷ yields to S3

79.6–87.5 **S3** [turn] Yeah, I was just gonna add, the psychological side is so underrated. Like, that's what separates good from great, honestly. ▷ yields, end

82.4–83.0 **S2** [backchannel] right

83.3–84.1 **S4** [backchannel] absolutely

test split · tone: professional and business-like · noise: cafe, 37.7 dB SNR; alarm beep · voices: S1 M, S2 F, S3 F, S4 F, S5 M
EOT 9 yields/2 holds · INT 1 interruptions/8 negatives (7 backchannels, 1 laughter) · 4 failed claims (excluded)

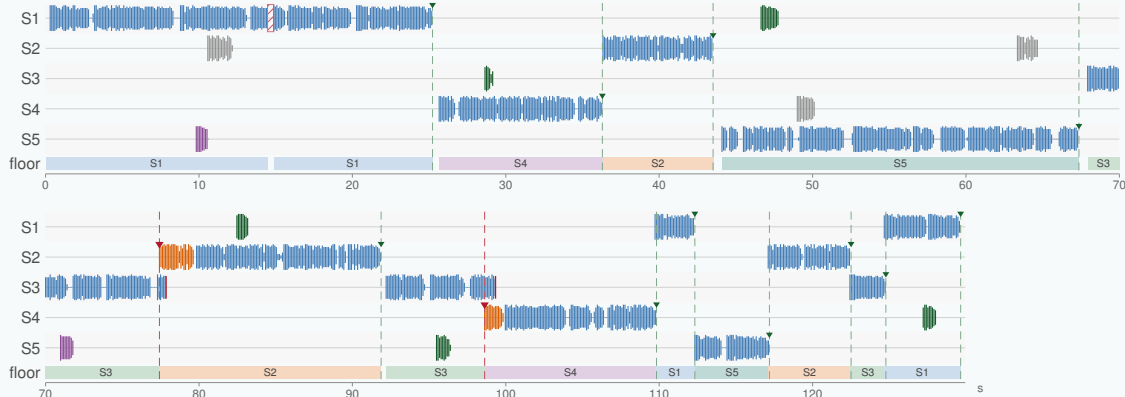


Transcript of test_00013

start—end (s) speaker [ground-truth label] text ▷ outcome

- 0.3–19.2 **S1** [turn] So, uh, I'll just start—last weekend I finally watched The Last Black Man in San Francisco. I know I'm late to the party, but honestly, it hit me harder than I expected. The cinematography is stunning, sure, but I was just not prepared for all that raw vulnerability. It actually reminded me a lot of when I read Colson Whitehead's Nickel Boys, how it just lingers with you for days. ▷ yields to S2
- 8.0–9.2 **S2** [backchannel] Oh, absolutely.
- 9.4–9.9 **S3** [backchannel] mm-hm
- 14.9–16.2 **S4** [failed claim] Yeah, that's a good one.
- 19.2–44.3 **S2** [turn] So, I haven't seen that—um—movie, but I did read The Nickel Boys. And, you know, that book just left me rattled. I sat on my couch for a bit after finishing it, not really sure what to do next. I think what I'm drawn to in stories lately is just, like, how they handle injustice without feeling sensational. I actually wanted to recommend This Is How You Lose the Time War. It's different, obviously—sci-fi, lyrical, but those letter exchanges? They're so personal, it almost reads like a memoir. ▷ holds, pauses
- 34.7–35.4 **S3** [failed claim] Right, I—
- 45.3–52.0 **S2** [turn] —haven't heard of anyone who wasn't moved by that book. Even my friend who dislikes sci-fi ended up reading it twice. ▷ yields to S3
- 51.9–77.8 **S3** [turn] Yeah, exactly. I actually have a hard time with books that are, you know, just sad for the sake of being sad. That one wasn't like that though. But okay, so if you want something that's like, dark but also sort of funny in a twisted way, can I recommend the film In Bruges? It's about two Irish hitmen lying low after a job goes wrong. Super bleak, but I found myself laughing at lines I definitely shouldn't have laughed at. Maybe that says something about me. ▷ cut off
- 77.0–81.4 **S5** [interruption] I, uh—can I just say, In Bruges is one of those movies that I, um, keep meaning to watch but— ▷ takes floor from S3 ▷ holds, pauses
- 81.8–89.5 **S5** [turn] I sort of avoid it every time because I assume it's just going to be too dark. But maybe I've been missing the point if there's—like—a real humor thread there? ▷ yields to S3
- 89.5–101.4 **S3** [turn] Oh, totally. It's miserable and ridiculous at the same time. Like, existential slapstick. You've got to be in the right mood, though. ▷ yields to S4
- 101.9–118.5 **S4** [turn] Yeah, speaking of mood, I'm going to take us in a different direction. I recommend Rules of Civility by Amor Towles. Whenever I hit a reading slump, that's the one I go back to. It's all sharp dialogue and this jazzy, 1930s New York atmosphere. Not a lot of existential dread, just... clever people trying to outsmart their lives.
- 106.6–107.8 **S2** [backchannel] oh interesting
- 112.2–113.1 **S5** [failed claim] I—hang on—
- 113.4–115.1 **S1** [failed claim] That's been on my list for ages.
- 118.6–123.8 **S4** [turn] —And it's got this sort of bittersweet retro optimism, so, not the kind of book that'll wreck your week, promise. ▷ yields to S5
- 124.3–141.9 **S5** [turn] Okay, so, uh, on a totally different note, if you're interested in documentaries, I'd suggest American Factory. It's not easy viewing, but it's exactly the kind of thing that... well, I kept pausing it just to process. A lot about, you know, hope versus reality, and how work changes us. Even my dad, who usually hates documentaries, sat down and finished it with me. ▷ yields to S1
- 142.0–145.1 **S1** [turn] Yeah, I saw that last year. It was sobering, but so well done. ▷ yields to S2
- 142.8–143.8 **S2** [backchannel] Definitely.
- 143.4–143.9 **S3** [backchannel] right
- 144.5–145.1 **S4** [backchannel] Uh-huh.
- 145.5–151.7 **S2** [turn] It sounds like most of us go for the heavy stuff—maybe next time someone should bring a comedy, just for balance. ▷ yields to S3
- 151.7–157.2 **S3** [turn] Haha, yeah, we're really out here trying to see who can cry the hardest. ▷ yields, end
- 154.6–155.5 **S1** [laughter] haha right
- 155.6–156.1 **S5** [backchannel] yeah

test split · tone: enthusiastic and energetic · noise: outdoor park, 25.2 dB SNR; dishes clinking, door knock · voices: S1 F, S2 F, S3 M, S4 M, S5 F
EOT 11 yields / 1 holds · INT 2 interruptions / 7 negatives (5 backchannels, 2 laughter) · 3 failed claims (excluded)



Transcript of test_00023

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–14.5 **S1** [turn] Hi everyone! So, I wanted to kick off with a quick story—last week, I tried to schedule my annual checkup and, oh my goodness, it took me almost forty minutes on the phone. I kept getting put on hold and then the receptionist thought I wanted a dental cleaning, not a physical. It was almost comical. ▷ holds, pauses

9.8–10.6 **S5** [laughter] Haha, oh no!

10.6–12.2 **S2** [failed claim] That sounds so stressful.

14.9–25.2 **S1** [turn] It really was! And then, when I finally got it sorted, they only had appointments at seven thirty in the morning. I'm not even awake before eight most days, so, um, now I have to go to bed at like nine the night before. ▷ yields to S4

25.7–36.3 **S4** [turn] Yeah, uh, I get that. I tried to get a cardiologist appointment last month and, well, I just kept putting it off, because every time I checked, the website said 'call for availability,' and I... I just never called. ▷ yields to S2

28.6–29.2 **S3** [backchannel] Mm-hm.

36.3–43.5 **S2** [turn] I always wish there was just one button that booked everything at once, like, you know, for your yearly stuff, flu shot, dentist, whatever. ▷ yields to S5

44.1–56.5 **S5** [turn] Oh! Totally! Last year, I used that new patient portal app, right, and it let me see every single available time slot. I felt like I was winning the lottery when I got an afternoon spot on a Friday.

46.6–47.8 **S1** [backchannel] That's amazing!

49.0–50.2 **S4** [failed claim] Whoa, how did you—

56.6–67.4 **S5** [turn] —Haha, seriously! But then I told my mom about it and she somehow managed to sign herself up for a kid's appointment for vaccinations instead of her physical. She called me in a panic— ▷ yields to S3

63.4–64.7 **S2** [failed claim] Oh no, poor thing.

67.9–77.9 **S3** [turn] I had a similar thing, um, when I tried booking a specialist for my ankle. The online calendar said 'book now,' but then the confirmation email was all in medical codes. I spent twenty minutes just Googling what I'd signed up for. ▷ cut off

71.0–71.8 **S5** [laughter] Haha, been there!

77.4–79.7 **S2** [interruption] Did you end up with the right appointment, though? ▷ takes floor from S3

79.8–91.9 **S2** [turn] Because I once accidentally booked, uh, an oncology consult when I meant to get dermatology, and I only realized the mix-up when they called me, super concerned, to check if I'd gotten a cancer diagnosis. ▷ yields to S3

82.5–83.3 **S1** [backchannel] Oh wow!

92.2–99.3 **S3** [turn] No, I, uh, had to call them and wait on hold forever. But the nurse was super friendly once I got through. She even rescheduled me for the following week, and, um, walked me through each of the appointment codes. ▷ cut off

95.5–96.4 **S5** [backchannel] That's so nice!

98.6–99.9 **S4** [interruption] Actually, can I— ▷ takes floor from S3

100.0–109.8 **S4** [turn] Uh, I once showed up at the wrong location because I didn't realize the clinic had moved. I walked in and it was just...a bakery. The staff there gave me a cupcake, though, so it wasn't a total loss. ▷ yields to S1

109.7–112.3 **S1** [turn] That's the best scheduling reward ever, honestly. ▷ yields to S5

112.3–117.2 **S5** [turn] Haha, scheduling with snacks! I'd show up anywhere if you promised me a cupcake at the end. ▷ yields to S2

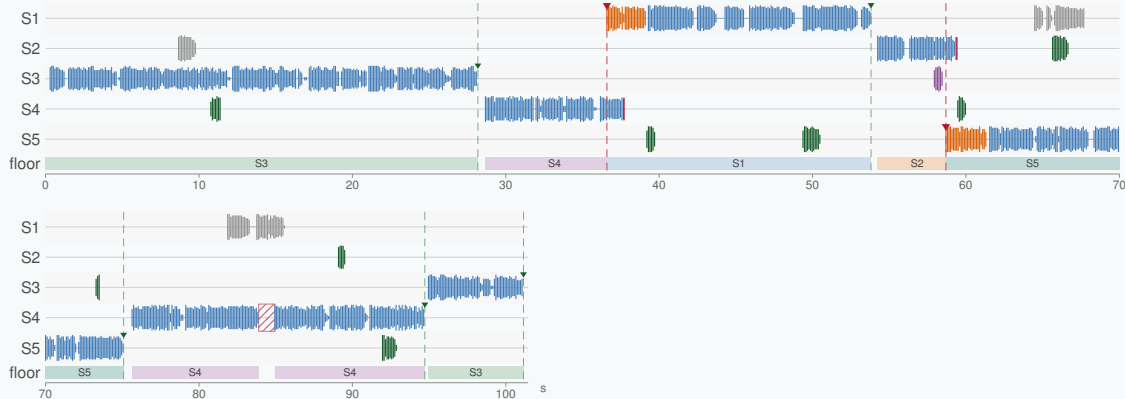
117.1–122.5 **S2** [turn] Right! Maybe that's what clinics are missing—it should be like getting your haircut, where you get a little treat. ▷ yields to S3

122.4–124.8 **S3** [turn] That would make waiting on hold a bit sweeter, yeah. ▷ yields to S1

124.7–129.7 **S1** [turn] You know, if I could schedule online and then pick up a muffin, I'd sign up for every screening under the sun. ▷ yields, end

127.2–128.1 **S4** [backchannel] Same here.

test split · tone: curious and inquisitive · noise: school hallway, 32.5 dB SNR; laughter offmic · voices: S1 F, S2 M, S3 M, S4 F, S5 F
EOT 5 yields / 1 holds · INT 2 interruptions / 9 negatives (8 backchannels, 1 laughter) · 3 failed claims (excluded)



Transcript of test_00038

start-end (s) speaker [ground-truth label] text > outcome

0.3–21.0 **S3** [turn] Morning, everyone. Before we deep dive, I just want to share something that, uh, happened yesterday. I was using our app at that old coffee shop across from the station—you know the one—and the owner actually recognized the interface and started asking me questions. Turns out, she really appreciated the new order summary screen, but was struggling with the feedback button. That kind of firsthand reaction, honestly, gave me a lot to think about for today.

8.7–9.8 **S2** [failed claim] That's interesting—

10.8–11.5 **S4** [backchannel] oh wow

21.1–28.2 **S3** [turn] —so, I guess my question for the group is whether we can make that feedback entry point more intuitive, maybe even context-specific? > yields to S4

28.7–37.7 **S4** [turn] Yeah, that actually reminds me of my grandma trying to use the app, which—she's usually pretty savvy, but she kept missing that button too. I think maybe we do need to rethink its placement or, like, the color contrast? I'm curious if anyone's experimented with alternatives. > cut off

36.6–39.1 **S1** [interruption] I actually ran a quick prototype test last week— > takes floor from S4

39.2–39.8 **S5** [backchannel] mm-hm

39.3–53.8 **S1** [turn] —where the feedback button animated in after a transaction, sort of drawing the eye. People responded more, but some said the animation felt slightly intrusive. It raises that tricky balance between visibility and, um, subtlety. > yields to S2

49.4–50.5 **S5** [backchannel] yeah exactly

54.2–59.4 **S2** [turn] I've noticed that with a few apps lately. Sometimes, like, those animated prompts make me feel rushed, or almost pressured? But—I guess if it's gentle enough... > cut off

57.9–58.5 **S3** [laughter] haha

58.7–61.4 **S5** [interruption] I actually, um, have a short story too, if I can— > takes floor from S2

59.5–60.0 **S4** [backchannel] right

61.5–70.6 **S5** [turn] —just to add, a friend of mine works in a bakery and she said, the first time she saw the feedback pop up, she thought it was a receipt question. Maybe we need a little icon or microcopy?

64.5–67.7 **S1** [failed claim] What if we, uh, put it in the order summary itself, like—

65.6–66.7 **S2** [backchannel] oh interesting

70.8–75.1 **S5** [turn] —that's a good thought, yeah. I think it could feel more like a conversation then, not a demand. > yields to S4

73.3–73.6 **S3** [backchannel] mm

75.6–83.9 **S4** [turn] So do we want to try a version where it's there but, like, ambient? Just an icon in the summary, and maybe if you haven't used it in a while, it gets a little nudge? > holds, pauses

81.9–85.6 **S1** [failed claim] That sounds reasonable to me, but maybe we need a quick user test—

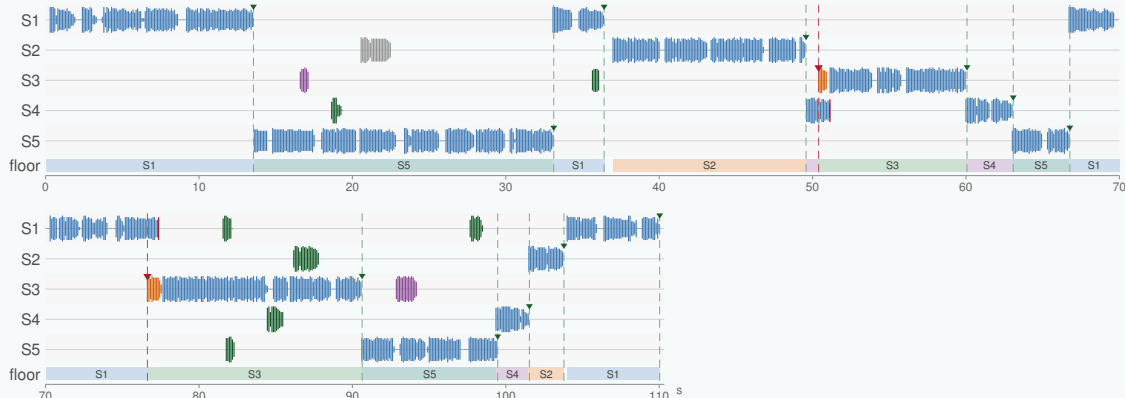
85.0–94.7 **S4** [turn] —or perhaps we just put that in this week's sprint as an experiment? I love the storytelling today, by the way. It feels like we're really close to something that helps actual people. > yields to S3

89.1–89.6 **S2** [backchannel] yeah

92.0–92.9 **S5** [backchannel] absolutely

95.0–101.2 **S3** [turn] Same here—just that little moment of seeing someone light up using the app, I mean, it's a reminder why this work matters. > yields, end

test split · tone: enthusiastic and energetic · noise: dorm room, 27.9 dB SNR · voices: S1 M, S2 M, S3 F, S4 F, S5 F
EOT 12 yields / 0 holds · INT 2 interruptions / 9 negatives (7 backchannels, 2 laughter) · 1 failed claims (excluded)



Transcript of test_00104

start-end (s) speaker [ground-truth label] text > outcome

0.3–13.6 **S1** [turn] Alright, good morning everyone. Let's kick off. I thought we'd start with our check-ins, but really, I want to hear some stories from this week. Anyone have a fun product moment to share before we dive into updates? > yields to S5

13.6–28.0 **S5** [turn] Oh, pick me! I had the weirdest user testing session yesterday. This guy literally tried to swipe on the logo, like he thought it was a magical button or something. Uh, he just kept poking it and laughing. It honestly made my whole afternoon.

16.6–17.2 **S3** [laughter] haha

18.6–19.3 **S4** [backchannel] Right.

20.6–22.5 **S2** [failed claim] That, um, actually reminds me—

28.1–33.1 **S5** [turn] —He said, 'is this the Easter egg?' Like, I guess we need a secret feature just for him now. > yields to S1

33.1–36.4 **S1** [turn] Haha, that's amazing. User logic is undefeated. > yields to S2

35.7–36.1 **S3** [backchannel] mm-hm

37.0–49.6 **S2** [turn] Um, so, actually, something similar happened on my feedback call. The user got confused by the microcopy on the onboarding. She kept asking if she had to connect her bank, which isn't even a thing on our app. I, uh, realized we might've been too... clever? > yields to S4

49.6–51.1 **S4** [turn] I flagged that last week. > cut off

50.4–51.0 **S3** [interruption] Wait— > takes floor from S4

51.1–60.1 **S3** [turn] Yeah, I just saw the latest build and still see the same line. Can we swap it today? It's muddy. Like, users keep DMing support about banking and, honestly, it's wasting our time. > yields to S4

60.0–63.1 **S4** [turn] Yes. I can update the string. I'll push it before lunch. > yields to S5

63.0–66.8 **S5** [turn] Love it! Can we make it sparkle or something? Kidding. I mean, unless? > yields to S1

66.7–77.4 **S1** [turn] Haha, a glitter button on onboarding, it's very you. Okay, direct action item—copy tweak, no confetti. Um, anyone else want to share a mini-disaster or a hero moment before we do blockers? > cut off

76.6–77.6 **S3** [interruption] I've got one— > takes floor from S1

77.7–90.6 **S3** [turn] Yesterday I was heads down prototyping and accidentally sent a push notification to our staging testers at 2 a.m. My phone blew up. Nobody was mad, but I woke up to five memes about sleep deprivation. Honestly, it was kind of wholesome. > yields to S5

81.6–82.2 **S1** [backchannel] Classic.

81.8–82.3 **S5** [backchannel] No way!

84.5–85.5 **S4** [backchannel] Good bonding.

86.2–87.9 **S2** [backchannel] Sounds oddly comforting.

90.6–99.5 **S5** [turn] Okay but, like, we've all done way worse. Remember the time I, um, demoed the wrong build in front of the CEO? The one with the rainbow cats background? > yields to S4

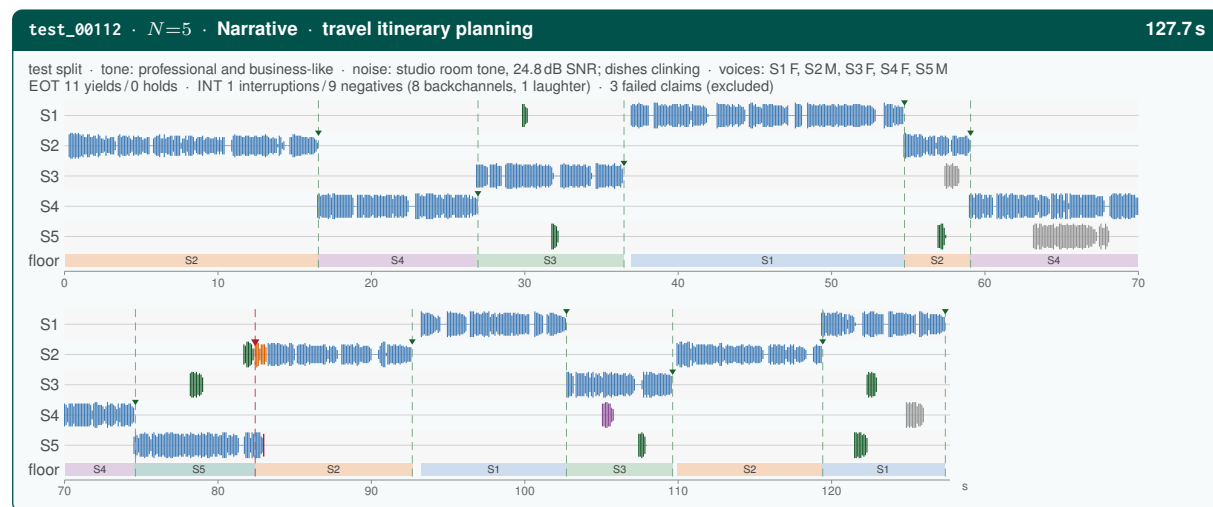
92.9–94.3 **S3** [laughter] Haha, never forget.

97.7–98.5 **S1** [backchannel] Legendary.

99.4–101.5 **S4** [turn] We should ship it as an April Fool's. > yields to S2

101.5–103.8 **S2** [turn] That, um, actually might boost engagement. > yields to S1

104.0–110.0 **S1** [turn] Okay, I love the energy today. Let's take this momentum into blockers. Who wants to go first? > yields, end



Transcript of test_00112

start-end (s) speaker [ground-truth label] text > outcome

0.3–16.6 **S2** [turn] Alright, so I thought it might be helpful if I start with a quick story—last time I tried to plan a multi-city trip, I ended up stuck in the Zurich airport for six hours because I booked the connections way too tight. So this time, I figured we should build in more buffer between our flights. Uh, does that sound okay to everyone? > yields to S4

16.5–27.0 **S4** [turn] Absolutely, that sounds like a great idea! I missed a train in France once because of a ten-minute delay. Honestly, travel always throws curveballs, so I love the buffer plan. > yields to S3

26.9–36.5 **S3** [turn] Yeah, same—oh my gosh, remember when I got stranded in Milan on that semester abroad? You have to leave breathing room or, like, everything immediately falls apart. > yields to S1

29.9–30.3 **S1** [backchannel] Mm-hm.

31.8–32.2 **S5** [backchannel] Yeah.

36.9–54.8 **S1** [turn] So, if we're all agreed, I'd like to suggest we prioritize direct flights where possible. I know it sometimes costs more, but honestly, the risk of missing connections—I mean, I once had to rebook an entire leg just because a thirty-minute delay snowballed. It's not worth the stress. > yields to S2

54.7–59.1 **S2** [turn] Right, and if we do have layovers, maybe aim for, let's say, no less than two hours? > yields to S4

56.9–57.5 **S5** [backchannel] Uh-huh.

57.4–58.3 **S3** [failed claim] Sounds good to me.

59.0–64.3 **S4** [turn] Yeah, two hours feels safe—especially with international terminals and all those security lines.

63.2–68.1 **S5** [failed claim] Can I, um, maybe suggest we also look at hotels near the train stations? Like, uh—

64.4–74.6 **S4** [turn] —oh, good point! That definitely makes early departures easier. I had a night once where I thought I could just take a cab but there weren't any running at 4am, so yeah, proximity matters. > yields to S5

74.5–83.0 **S5** [turn] Right, yeah, I just—last time in Vienna, I almost missed my train because my hotel was way too far from the station. So I'd feel better if we could, uh, keep that in mind. > cut off

78.2–79.1 **S3** [backchannel] Totally agree!

81.7–82.3 **S2** [backchannel] For sure.

82.4–83.2 **S2** [interruption] Actually— **> takes floor from S5**

83.3–92.7 **S2** [turn] —I was just going to say, when we're in Prague, the main train station area has a bunch of hotels, but I think they book up fast around August. Should I lock that in sooner rather than later? **> yields to S1**

93.3–102.7 **S1** [turn] I think that would be wise. The earlier we reserve, the less likely we end up paying a premium for the last available rooms—or worse, split up across town. **> yields to S3**

102.7–109.6 **S3** [turn] Haha, I don't want a repeat of Budapest where I had to sleep above a karaoke bar. I still have nightmares about that night. **> yields to S2**

105.1–105.8 **S4** [laughter] Haha, right.

107.4–107.9 **S5** [backchannel] Oh no.

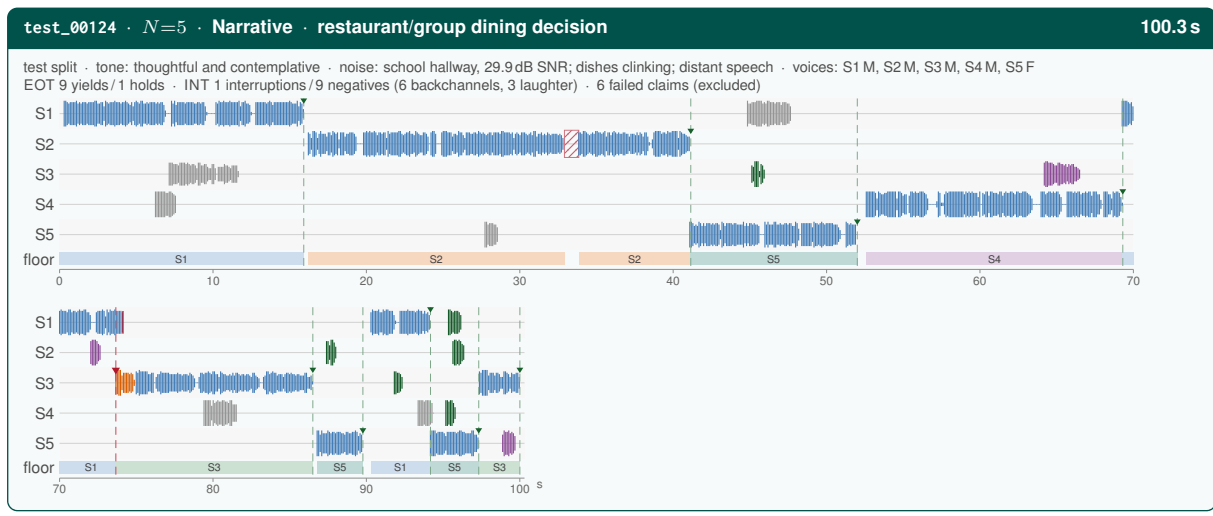
109.9–119.4 **S2** [turn] Okay, so I'll draft up a preliminary itinerary with relaxed connections and those centrally located hotels, and we can review as a group before booking anything. Does that work? **> yields to S1**

119.3–127.4 **S1** [turn] Yes, that sounds like a good path forward. Once you have that document, I'll review for any possible gaps, but I think the approach is solid. **> yields, end**

121.5–122.4 **S5** [backchannel] Yeah, thank you.

122.3–123.0 **S3** [backchannel] Perfect!

124.9–126.0 **S4** [failed claim] I'm good with that too.



Transcript of test_00124

start–end (s) speaker [ground-truth label] text **>** outcome

0.3–15.9 **S1** [turn] You know, choosing a restaurant always takes me back to when my family used to spend almost an hour just driving around town debating where to go, and half the time we'd just end up at the same diner. There's something oddly comforting about that whole ritual, even if it meant we were all starving by the time we actually sat down. **> yields to S2**

6.2–7.6 **S4** [failed claim] Yeah, that sounds familiar.

7.1–11.7 **S3** [failed claim] Same here! We'd always argue about Italian versus Thai until someone just gave up.

16.2–32.9 **S2** [turn] I, uh, actually kind of dread group dinners because of that. Like, I remember one time, my friends and I were trying to pick a brunch spot in Seattle, and I—I offered this place I liked but then, you know, I worried too much that maybe everyone else would hate it, so I just sort of stopped suggesting things after a while. **> holds, pauses**

27.7–28.6 **S5** [failed claim] Oh come on—

33.8–41.1 **S2** [turn] —and, uh, we ended up at the first place someone else mentioned, which was just okay. So, I don't know, I guess I'm not great at this. **> yields to S5**

41.1–52.0 **S5** [turn] But honestly, someone has to make the call, you know? I don't think anybody really cares that much half the time. Like, last week, I just picked a taco place and said, 'We're meeting there at seven,' and everyone went along with it. No drama. **> yields to S4**

44.8–47.7 **S1** [failed claim] It's true, sometimes just having a decision is a relief.

45.1–46.0 **S3** [backchannel] Exactly.

52.6–69.3 **S4** [turn] You know, my own approach is just to ask about allergies and then go for variety. Like, one time I organized a company dinner and picked this Mediterranean spot thinking there'd be something for everyone. It actually went pretty well, except for one person who didn't like olives, but that was pretty minor. **> yields to S1**

64.2–66.5 **S3** [laughter] Haha, yeah, there's always that one person...

69.3–74.1 **S1** [turn] I had a friend who brought her own hot sauce to every restaurant, so, uh, I think there's always going to be that wildcard. **> cut off**

72.0–72.7 **S2** [laughter] Haha right.

73.7–74.9 **S3** [interruption] Oh, actually, I, uh— **> takes floor from S1**

75.0–86.5 **S3** [turn] —I totally relate to that! My cousin refuses to eat anything that isn't beige. So, if it's not bread or pasta, he just picks at it and then orders fries. It's kind of hilarious but makes planning dinners nearly impossible. **> yields to S5**

79.4–81.6 **S4** [failed claim] Oh man, picky eaters are a real challenge.

86.8–89.8 **S5** [turn] That's why I just ask what people won't eat and ignore the rest. **> yields to S1**

87.4–88.1 **S2** [backchannel] Makes sense.

90.3–94.2 **S1** [turn] So, do we want to try that tonight? Maybe just pick a place and not overthink it? **> yields to S5**

91.8–92.4 **S3** [backchannel] I'm in.

93.4–94.4 **S4** [failed claim] That works for me.

94.2–97.3 **S5** [turn] Alright, I'll make a call. No one's allergic to anything weird, right? **> yields to S3**

95.2–95.8 **S4** [backchannel] Me too.

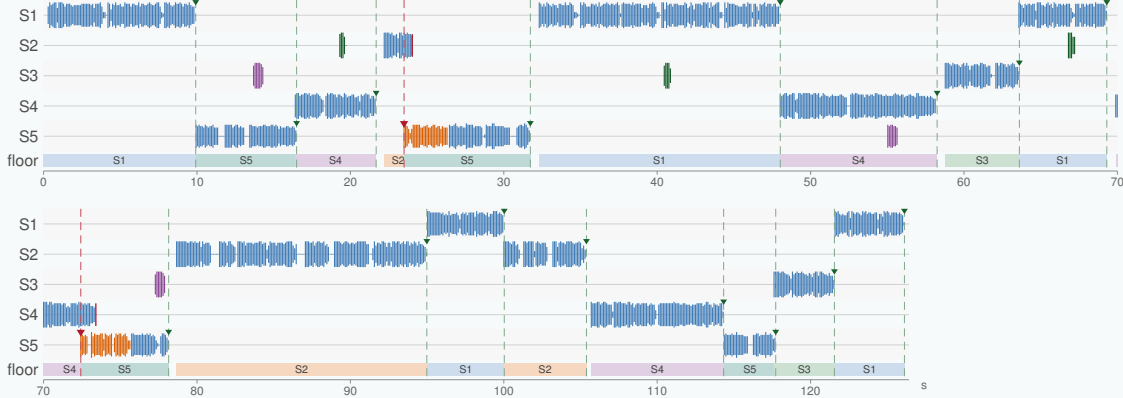
95.4–96.2 **S1** [backchannel] All clear.

95.6–96.4 **S2** [backchannel] Nope, I'm good.

97.4–100.0 **S3** [turn] Just no olives for me, please—sorry, couldn't resist. **> yields, end**

98.9–99.8 **S5** [laughter] Ha! Noted.

test split · tone: curious and inquisitive · noise: cafe, 20.3 dB SNR; school bell; distant speech · voices: S1 F, S2 M, S3 F, S4 M, S5 F
EOT 16 yields / 0 holds · INT 2 interruptions / 6 negatives (3 backchannels, 3 laughter) · 0 failed claims (excluded)



Transcript of test_00129

start-end (s) speaker [ground-truth label] text ▷ outcome

0.3–9.9 **S1** [turn] So, um, has anyone actually decided what they want for dinner? I feel like we could spend the whole night just debating menus again, like last time at that Thai place. ▷ yields to S5

9.9–16.5 **S5** [turn] Oh yeah, that was chaos! Remember when we spent, like, thirty minutes just arguing about whether curry counts as a soup? ▷ yields to S4

13.7–14.4 **S3** [laughter] haha

16.4–21.7 **S4** [turn] Honestly, I still say it does. And I still want something spicy tonight, just putting that out there. ▷ yields to S2

19.3–19.7 **S2** [backchannel] Mm-hm

22.2–24.0 **S2** [turn] But I really think we should try somewhere new this time— ▷ cut off

23.5–26.4 **S5** [interruption] —wait, what about that Ethiopian spot near my place? ▷ takes floor from S2

26.5–31.7 **S5** [turn] They do like, this incredible lentil stew, and you eat everything with your hands. Super fun. ▷ yields to S1

32.3–48.0 **S1** [turn] Oh, that sounds really cozy, actually. I tried Ethiopian once years ago with my cousin, and it was, I don't know, just really communal? Like everyone sharing from the same plate, dipping injera, laughing together—it felt like a totally different vibe than just ordering my own dish. ▷ yields to S4

40.5–40.9 **S3** [backchannel] Yes

48.0–58.2 **S4** [turn] I can get behind that. Plus, I love when you don't have to fuss with forks and knives. Last time I went, I spilled water everywhere, but at least I didn't launch a utensil across the room for once. ▷ yields to S3

55.0–55.7 **S5** [laughter] haha right

58.8–63.6 **S3** [turn] I always get nervous about getting food all over my hands though. Does it get, like, super messy? ▷ yields to S1

63.6–69.3 **S1** [turn] Not if you're careful! And honestly, I think that's part of the charm. You just kind of lean in and go with it. ▷ yields to S4

66.8–67.3 **S2** [backchannel] Right

69.9–73.4 **S4** [turn] Yeah, but what if someone double dips? You know there's always one person— ▷ cut off

72.4–75.7 **S5** [interruption] —Guilty! I totally do that if there's, like, good sauce left. ▷ takes floor from S4

75.7–78.2 **S5** [turn] But I promise, I'll restrain myself. Maybe. ▷ yields to S2

77.3–78.0 **S3** [laughter] Haha

78.7–95.0 **S2** [turn] Okay, okay, but can I throw in a quick story? When I was in college, my roommate took me to a Persian restaurant and insisted we eat with our hands. I'd never felt more out of place—everyone was just going for it and there I was, awkwardly trying to peel bread apart without dropping stuff everywhere. ▷ yields to S1

95.0–100.0 **S1** [turn] Oh wow, I can totally picture you just quietly panicking but pretending you're cool about it. ▷ yields to S2

100.0–105.4 **S2** [turn] Haha, exactly! But it kind of broke the ice, you know? After a few minutes, everyone just relaxed. ▷ yields to S4

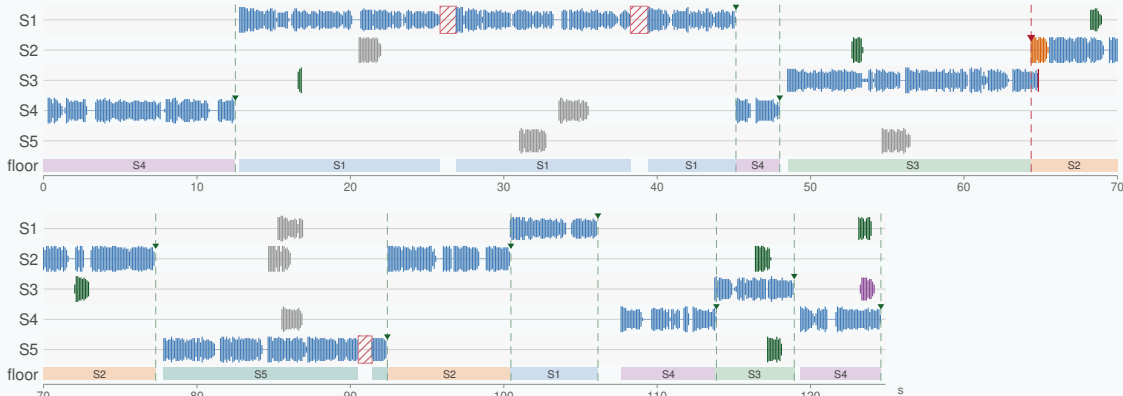
105.7–114.3 **S4** [turn] That's the best, honestly. Food kind of does that, just pulls everyone together. So, are we settled on Ethiopian, or should I pitch the taco truck on 9th? ▷ yields to S5

114.4–117.7 **S5** [turn] Nooo, now I'm stuck on lentils. Ethiopian's got my vote. ▷ yields to S3

117.6–121.6 **S3** [turn] I'm good with that too. Let's call ahead just in case—it gets packed on weekends. ▷ yields to S1

121.6–126.1 **S1** [turn] Alright, Ethiopian it is, but if it turns into a food fight, I'm blaming you all. ▷ yields, end

test split · tone: thoughtful and contemplative · noise: dorm room, 31.4 dB SNR; chairs scraping, keyboard clatter · voices: S1 M, S2 M, S3 M, S4 M, S5 F
EOT 10 yields / 3 holds · INT 1 interruptions / 8 negatives (7 backchannels, 1 laughter) · 7 failed claims (excluded)



Transcript of test_00171

start—end (s) speaker [ground-truth label] text ▷ outcome

0.3–12.5 **S4** [turn] Alright everyone, thanks for joining today. We're here to talk about people's experiences with scheduling medical appointments, and I'd love if we could start with a story or two. Who'd like to kick us off? ▷ yields to S1

12.8–25.8 **S1** [turn] Oh I've got one! So last month, I needed to book a follow-up for my knee, right? I thought it'd be simple—you know, just call in, give 'em my details, pick a slot. But, uh, I ended up playing phone tag for, I kid you not, a full week before anyone actually got back to me. ▷ holds, pauses

16.6–16.9 **S3** [backchannel] mm

20.6–22.0 **S2** [failed claim] Yeah, that sounds about right.

26.9–38.3 **S1** [turn] Yeah exactly, and after all that, when I finally, you know, got a hold of someone, she told me the earliest they could do was three weeks out. And I, uh, I had to haggle just to get on a waitlist for cancellations. It's wild. ▷ holds, pauses

31.0–32.8 **S5** [failed claim] That's super common, honestly.

33.6–35.5 **S4** [failed claim] Was that your primary care office or—

39.4–45.1 **S1** [turn] —no, it was the ortho clinic. My PCP is actually pretty good, but specialists, you know, totally different pace. ▷ yields to S4

45.2–48.0 **S4** [turn] Right, makes sense. Thanks for sharing that, Alex. ▷ yields to S3

48.5–60.3 **S3** [turn] I remember when it was just you and, uh, a landline, and the receptionist wrote everything down by hand. Now, it's all these portals and forms. Last time, I spent half an hour trying to reset a password just to see an open slot.

52.7–53.4 **S2** [backchannel] Same here.

54.7–56.6 **S5** [failed claim] Oh, those portals drive me nuts.

60.4–64.8 **S3** [turn] And after all that effort, by the time I got logged in? All the appointments I wanted were gone. It's like trying to buy concert tickets these days—blink and you miss it. ▷ cut off

64.4–65.5 **S2** [interruption] Let me just jump in— ▷ takes floor from S3

65.6–77.3 **S2** [turn] —I actually had an odd experience with their automated call-back system. I scheduled, got a confirmation, but then... no call. Turns out it never properly logged in their system, but there was no way for me to know. ▷ yields to S5

68.3–69.0 **S1** [backchannel] Oh man.

72.0–73.0 **S3** [backchannel] That's the worst.

77.8–90.5 **S5** [turn] Honestly, I just go down there in person when I'm sick of phone trees. It's not efficient but at least I know I'll get a real answer. The last time, I waited in the lobby for twenty minutes, and the guy ahead of me had been on hold for forty-five minutes before giving up. ▷ holds, pauses

84.7–86.1 **S2** [failed claim] Can I—wait, quick question—

85.3–87.0 **S1** [failed claim] Sometimes old school works best!

85.5–86.9 **S4** [failed claim] Wow, you're persistent, Claire.

91.4–92.4 **S5** [turn] —yeah go ahead. ▷ yields to S2

92.5–100.5 **S2** [turn] Did anyone here ever try those text reminders they offer? I signed up, but honestly, I got way too many. It was like spam after a while. ▷ yields to S1

100.4–106.1 **S1** [turn] Same! I got a text, a call, and then they emailed me about the same appointment. I appreciate the effort but wow. ▷ yields to S4

107.6–113.9 **S4** [turn] Alright, we're nearly at time. Does anyone want to share a quick tip or, uh, lesson learned before we wrap? ▷ yields to S3

113.8–118.9 **S3** [turn] Just patience, really. And if you get a human, be extra nice—you never know when they'll squeeze you in. ▷ yields to S4

116.4–117.4 **S2** [backchannel] Yeah, definitely.

117.2–118.2 **S5** [backchannel] Amen to that.

119.4–124.6 **S4** [turn] Thanks everyone, this was great. Let's hope next time scheduling'll be a bit more humane. ▷ yields, end

123.1–124.0 **S1** [backchannel] Fingers crossed!

123.2–124.2 **S3** [laughter] haha, yeah
