Are you laughing at me? Patterns of laughter while playing a voice activated game in pairs compared to single-player

Talha Bedir
University of Gothenburg
talha.bedir@gu.se

Bora Kara
University of Gothenburg
guskarabo@student.gu.se

Christine Howes
University of Gothenburg
christine.howes@gu.se

Abstract

Voice-activated agents, such as Alexa, are now commonplace, but little is known about how interactions with such agents are influenced by whether there is another human co-present in the interaction or not. We present a pilot study of people playing a voice-activated game in either a 'solo' or 'duo' condition. Preliminary analyses suggest that laughter is more common in the duo condition, but triggered in specific contexts since the system is not treated as a full dialogue participant, with participants laughing together at the system when it produces errors, for example.

1 Introduction

Interaction with voice-activated agents, such as Alexa or Siri, is now commonplace, but little is known about the differences in interactions when they do or do not include another human participant. Research suggests that people are more creative when in dialogue with another person rather than performing a task on their own (Howes et al., 2015), and that laughter is more common in social situations (Provine, 2004). But do people treat voice systems as active participants in a dialogue?

Prior work additionally shows that people laugh to smooth breakdowns (Glenn, 2003) and slip into "careful speech" after misrecognitions (Oviatt et al., 1998a). We are therefore interested in whether and how these behaviours are different when interacting with a voice-activated system in a solo versus cooperative setting.

We present a pilot experiment using a voice-activated game based on an existing point-and-click thriller called Dr. Stanley's House.

2 Background

2.1 Laughter

Perhaps counterintuitively, only around one tenth of laughs are triggered by jokes (Provine, 2001), with most laughs serving dialogue functions such as cohesion and alignment (Mazzocconi et al., 2020; Koutsombogera and Vogel, 2022).

Additionally, laughing at something is much more likely if the person is laughing *with* somebody. Provine (2004) reports that (according to students' self-reports) laughter is 30 times more likely when another person is present.

Voice agents can also trigger laughter between humans, especially when the technology performs poorly, awkwardly or sometimes more smartly than expected Perkins Booker et al. (2024). These laughters are less frequent than human-human laughters and are often *at* the voice agent, rather than *with* it (Glenn, 2003).

Literature on video games further suggests that inter-player laughter plays a significant role in co-operation, teamwork and cohesion between players (Rychlowska et al., 2022) and strengthens group identity (Tekin, 2023).

We predict that players of a voice-controlled game will produce (i) affiliative co-laughter that reinforces teamwork, and (ii) laughters at the system's mistakes. If the system is being treated as a full dialogue participant, we would expect these to appear in both solo and duo interactions, whilst if not we should only see the first type in the duo condition.

2.2 Careful Speech

When a spoken interface misrecognises the input, speakers shift into a *careful-speech*, characterised by a slower rate, expanded vowel space, higher mean pitch and greater intensity (Oviatt et al., 1998b). Efficiently, it targets the troublesome segment first while also applying a mild global slowdown (Stent et al., 2008). Because careful speech reliably boosts automatic-speech-recognition (ASR) accuracy, it can be seen as an *interactional repair resource* that keeps the dialogue moving.

Proceedings of the 29th Workshop on the Semantics and Pragmatics of Dialogue, September 3–5, 2025, Bielefeld, Germany.

In multi-user households, the first repair attempt is often undertaken by a different family member who literally "speaks for Alexa," redistributing the burden of careful speech (Porcheron et al., 2018). A similar division of labour is found in interactions with a voice calendar: careful speech accounted for 40% of error-recovery overall, but was less frequent when a teammate stepped in to rephrase the request (Myers et al., 2018).

These patterns suggest a dual role for careful speech in voice gaming. Solitary players must rely on their own articulatory adjustments to maintain system intelligibility, whereas those in pairs can offload repair through turn-exchange. We hypothesise that solo play will induce more careful speech than cooperative play.

3 Methods and results

The game was developed using SpeechState[1], an open-source, browser-based spoken dialogue system built on the XState[2] library. Front-end development was facilitated using Vite,[3] and HTML was employed for integrating images, video, and sounds.

The visual and narrative elements of the game were adapted from *Dr. Stanley's House*, a freely available Flash-based puzzle game created by James Li (2005), in which players assume the role of a detective solving a mystery through interaction with the environment. Progress is controlled through conditional access to different "states" or rooms, which depend on the accumulation of specific items or information. In our adaptation the primary mode of interaction is voice.

Voice interaction was enabled using Microsoft Azure Speech Services, which provided speech-to-text conversion, text-to-speech (TTS) synthesis, and a natural language understanding (NLU) module, which enables interpretation of user input via detected intents and entities, which in turn triggers state transitions within the game.

For this pilot, we collected data from two sessions with two players playing together ("duo" condition) and two single-player sessions ("solo"), for a total of approximately 85 minutes of data (see table 1). For each session, we recorded participants' video, audio, and screen activity, and analysed the data for instances of laughter and careful speech.

[1]https://github.com/vladmaraev/speechstate
[2]https://github.com/statelyai/xstate
[3]https://github.com/vitejs/vite

Participant	Length	Laughter	CS
1	14	2	2
2	14	1	4
3	10	7	0
4	18	2	1
5, 6 pair	14	17, 9	2, 1
7, 8 pair	14	5, 6	1, 1

Table 1: Participant-wise laughter and careful speech (CS) tokens and play length

Qualitative analysis of laughs showed several triggers:

- Dramatic prosody or exaggerated scene descriptions
  - e.g., When the system says: "I enter the house. There is no voice besides my steps. Then suddenly someone called '(player's name)', my name!"
- Recognition failures
  - e.g., When the player commands an action and nothing happens on screen
- Unintended system responses (when players were talking to each other)
  - e.g., When the two players discuss "Should we go right?", the game takes it as an input
- Redundant scene descriptions, repeated when players issued a command multiple times
  - e.g., When the player(s) ask for help repeatedly in the same state
- Overlapping TTS outputs, when players moved through game states faster than the TTS could complete its output
  - e.g., When the player(s) issue moving command successively
- Unexpected silences, when the system paused longer than anticipated
  - e.g., Usually after recognition failures

Instances of careful speech were consistently observed following speech recognition failures by the system. These episodes were often accompanied by a change in body posture, such as leaning toward the microphone or screen, suggesting heightened attentiveness and a communicative repair effort.

Our pilot presents a promising strand of research. Our aim is to collect data from 10 groups per condition, to allow us to perform quantitative as well as qualitative analyses. The findings have implications for the status of the participant role of dialogue systems, which are simultaneously taken to be active participants in the dialogue whilst also being treated as distant.

Acknowledgements

This research was supported by ERC Starting Grant DivCon: Divergence and convergence in dialogue: The dynamic management of mismatches (101077927) and by the Swedish Research Council grant (VR project 2014-39) for the establishment of the Centre for Linguistic Theory and Studies in Probability (CLASP) at the University of Gothenburg.

References

Phillip Glenn. 2003. *Laughter in Interaction*. Studies in Interactional Sociolinguistics. Cambridge University Press.

Christine Howes, Patrick GT Healey, Pietro Panzarasa, and Thomas Hills. 2015. Ideas in dialogue: The effects of interaction on creative problem solving. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 37.

Maria Koutsombogera and Carl Vogel. 2022. Understanding laughter in dialog. *Cognitive Computation*, 14(4):1405–1420.

Chiara Mazzocconi, Ye Tian, and Jonathan Ginzburg. 2020. What's your laughter doing there? a taxonomy of the pragmatic functions of laughter. *IEEE Trans. on Affective Computing*, 13(3):1302–1321.

Chelsea Myers, Anushay Furqan, Jessica Nebolsky, Karina Caro, and Jichen Zhu. 2018. Patterns for how users overcome obstacles in voice user interfaces. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pages 1–7.

Sharon Oviatt, Jon Bernard, and Gina-Anne Levow. 1998a. Linguistic adaptations during spoken and multimodal error resolution. *Language and speech*, 41(3-4):419–442.

Sharon Oviatt, Margaret MacEachern, and Gina-Anne Levow. 1998b. Predicting hyperarticulate speech during human-computer error resolution. *Speech Communication*, 24(2):87–110.

Nynaeve Perkins Booker, Michelle Cohn, and Georgia Zellou. 2024. Linguistic patterning of laughter in human-socialbot interactions. *Frontiers in Communication*, 9:1346738.

Martin Porcheron, Joel E Fischer, Stuart Reeves, and Sarah Sharples. 2018. Voice interfaces in everyday life. In *proceedings of the 2018 CHI conference on human factors in computing systems*, pages 1–12.

Robert R Provine. 2001. *Laughter: A scientific investigation*. Penguin.

Robert R Provine. 2004. Laughing, tickling, and the evolution of speech and self. *Current Directions in Psychological Science*, 13(6):215–218.

Magdalena Rychlowska, Gary McKeown, Ian Sneddon, and William Curran. 2022. Laughter during cooperative and competitive games. In *Proceedings of the Workshop on Smiling and Laughter across Contexts and the Life-span within the 13th Language Resources and Evaluation Conference*, pages 32–34.

Amanda J Stent, Marie K Huffman, and Susan E Brennan. 2008. Adapting speaking after evidence of misrecognition: Local and global hyperarticulation. *Speech Communication*, 50(3):163–178.

Burak S Tekin. 2023. Cheering together: The interactional organization of choral vocalizations. *Language & Communication*, 88:73–89.