

HAI Alice - An Information-Providing Closed-Domain Dialog Corpus

Jelte van Waterschoot¹✉, Guillaume Dubuisson Duplessis², Lorenzo Gatti³,
Merijn Bruijnes¹ and Dirk Heylen¹

¹Human Media Interaction, University of Twente, 7522 NB, Enschede, The Netherlands

²CNRS, ISIR, Sorbonne Université, 75005, Paris, France

³FBK, University of Trento, 38123, Trento, Italy

l.gatti@fbk.eu, gdubuisson@telecom-paristech.fr

{j.b.vanwaterschoot, m.bruijnes, d.k.j.heylen}@utwente.nl

Abstract

The contribution of this paper is twofold: 1) we provide a public corpus for Human-Agent Interaction (where the agent is controlled by a Wizard of Oz) and 2) we show a study on verbal alignment in Human-Agent Interaction, to exemplify the corpus' use. In our recordings for the Human-Agent Interaction Alice-corpus (HAI Alice-corpus), participants talked to a wizarded agent, who provided them with information about the book *Alice in Wonderland* and its author. The wizard had immediate and almost full control over the agent's verbal and nonverbal behavior, as the wizard provided the agent's speech through his own voice and his facial expressions were directly copied onto the agent. The agent's hand gestures were controlled through a button interface. Data was collected to create a corpus with unexpected situations, such as misunderstandings, (accidental) false information, and interruptions. The HAI Alice-corpus consists of transcribed audio-video recordings of 15 conversations (more than 900 utterances) between users and the wizarded agent. As a use-case example, we measured the verbal alignment between the user and the agent. The paper contains information about the setup of the data collection, the unexpected situations and a description of our verbal alignment study.

Keywords: Corpus, Human-Agent Interaction, Wizard of Oz, Closed Domain, Information-Providing, Unexpected Situations, Verbal Alignment

1 Introduction

This paper presents the Human-Agent Interaction Alice-corpus (HAI Alice-corpus), a corpus of conversations between a user and an embodied conversational agent (ECA) operated by a wizard via the Wizard of Oz (WOz) method (Dahlbäck et al., 1993). The corpus has been collected as part of the ARIA-VALUSPA¹ project, in which a multi-modal virtual agent for information retrieval that can deal with unexpected situations is being developed. The goal of the project is to provide interested parties, such as fellow researchers and industry, with a toolkit for building their own virtual agent (Valstar et al., 2016; Bruijnes et al., 2013). We collected the corpus to investigate how users react to unexpected situations in a conversation with an autonomous state-of-the-art virtual agent. However, in a classic WOz approach where the wizard uses a button interface, it is nearly impossible to improvise in unexpected situations. This is why we gave our wizard the freedom to choose his own words and facial expressions, to create and respond to unexpected situations. To the best of our knowledge most state-of-the-art agents cannot yet cope with interruptions effectively, nor perform human-level verbal alignment (Dubuisson Duplessis et al., 2017b). However, we believe we are approaching such capabilities in agents, thanks to recent efforts in agent development. Therefore, we wanted to simulate an agent as closely as possible with this type of WOz setup. We analyzed the corpus for verbal alignment in Human-Agent (H-A) interaction. Verbal alignment is a process during a dialog where participants reuse lexical and syntactic structures (Pickering and Garrod, 2004). One example of reusing structures is by speaking with similar words in the conversation as other dialogue partici-



Figure 1: The conversational agent operated by the wizard

pants (Reitter et al., 2006).

First, we will describe some related corpora in Section 2. In Section 3 we will provide more details on the setup of our Wizard of Oz system. More details on the data collection will be given in Section 4. Section 5 will be a description of what will be released with the corpus, and an example of a dialog from the corpus. In Section 6 we briefly describe our use-case for the corpus on the topic of verbal alignment. In Section 7 we will discuss the limitations of the corpus, suggestions for its use, and some future studies.

2 Related Work

More and more dialog corpora are available involving Human-Human interactions as well as Human-Machine interactions; see (Serban et al., 2015) for a recent and extensive study. However, a closer look at the Human-Machine

¹www.aria-agent.eu

corpora reveals corpora are mostly available for spoken dialog systems, but very few are available in the Human-Agent community. The biggest difference between these corpora is that agents are often embodied and use multi-modal information, compared to only speech in spoken dialog systems. A well-known example of a corpus in H-A is the SEMAINE corpus (McKeown et al., 2010), dedicated to emotions in H-A Interactions. Recent efforts in collecting interaction data between a chatbot and a user include the UCAR corpus (Dubuisson Duplessis et al., 2016)², as well as the (RE-)WOCHAT international effort, which aims at collecting and annotating dialog data for chatbots and conversational agents (D’Haro et al., 2016)³.

We believe that more H-A corpora can be very useful to analyze H-A interaction, in order to improve agents’ communicative capabilities (e.g. a corpora study about adaptation and verbal alignment in H-A Interaction (Dubuisson Duplessis et al., 2017b)). This paper and the associated HAI Alice-corpus are a humble step towards sharing such valuable corpora for the H-A Interaction community.

3 Wizard of Oz System

The goal of the WOz system was to allow a user to have an interaction with an accomplice, who appears to the user as a virtual agent, and to record this interaction. Additionally, the user should believe (s)he is interacting with an advanced autonomous agent system⁴.

3.1 Wizard and Agent

The wizard and participant were located in different rooms. The wizard could see and hear the participant through a Skype video-conference connection. The agent was rendered in real-time on the wizard’s PC, and sent over the Skype connection to the participant. The wizard was trained to speak like a robot, with little pitch variation and a rhythm that is characteristic of generated speech. Additionally, using VoiceMeeter⁵ and Reaper⁶, his voice was passed through audio filters to make it sound more robotic. The facial animations of the agent were controlled using the wizard’s facial expressions. The Unity game engine, using the FacePlus plugin by Mixamo⁷, mapped the wizard’s expressions on to the face of the agent. The body postures and gestures of the agent were controlled by the wizard through a button interface. With these tools, our wizard was able to improvise in unexpected situations, which would not have been possible in a button interface for classic wizards.

3.2 Recording

The system allowed for multi-modal data collection on the interaction between the user and wizard. The audio-video

data in this corpus consists of the audio of both the wizard and the participant and the video of the agent.

4 Data Collection

In this section we will explain the scenario in the corpus, the variations and manipulations, information about the participants and the experiment protocol.

4.1 Domain

The conversational topic of our corpus is Alice’s Adventures in Wonderland (also known as Alice in Wonderland), the famous novel of Lewis Carroll. For the data collection experiment, a virtual agent controlled by a wizard acted as an expert on Alice’s Adventures in Wonderland (AiW). The agent was standing in a library and had the appearance of a middle-aged male intellectual, as can be seen in Figure 1. The wizard controlling the agent was well informed on the topic of AiW. The agent’s utterances were prepared beforehand, based on possible questions that the agent might be asked during the interaction. To prepare for possible questions, a survey was held before the experiment and a pilot study was conducted via Reddit. Additionally, agent utterances to suggest topics were generated. The wizard studied these utterances, so that he could respond as fast as possible. He was allowed to paraphrase the utterances, as long as the meaning would remain the same. The participant, uninformed of the WOz setup, was instructed to ask questions about AiW that the agent had knowledge about: the events of the book, its characters, the author Lewis Carroll, and some other information around the book, such as its adaptations in films and others media.

4.2 Experimental Variations

Two variations of the AiW scenario exist. Condition **A** was intended to prime the user to try out the agent’s capabilities rather than asking relevant questions on the book. Condition **B** was intended to be a credible version of an advanced and knowledgeable virtual agent that would increase the users’ curiosity on the book.

- Condition **A**: the agent was not knowledgeable on the book, often repeated the same sentences (“*I unfortunately don’t know about that*”), and intentional mistakes (such as using the wrong gender when saluting, or giving an incorrect answer to a question);
- Condition **B**, the agent answered questions to the best of its ability (and hid his ignorance when he could not answer), tried not to repeat himself, and took more time before answering.

4.3 Unexpected Situations

One aim was to record how participants respond to unexpected situations during an interaction with a virtual agent. We intentionally introduced sources of distraction and mistakes by the agent. These mistakes reflect technical limitations that exist in current state-of-the-art systems and that therefore often occur during a conversation with an autonomous agent. In particular, we created the following unexpected situations:

²Available at <https://ucar.limsi.fr>

³<http://workshop.colips.org/wochat/data/>

⁴For this reason we refer to the embodiment as an agent, and not as an avatar.

⁵<http://www.vb-audio.com/Voicemeeter/>

⁶<https://www.reaper.fm/>

⁷<https://www.mixamo.com>. Unfortunately, the FacePlus plugin has been removed from the Mixamo package since August 2017: <http://blogs.adobe.com/adobecare/2017/05/23/download-assets-from-mixamo/>

#	Sp.	Utterance
1	U	Hello
2	A	Hello madam, my name is James
3	U	I'm not a madam, but okay
4	A	I know everything about "Alice in Wonderland." Is there anything you would like to know?
5	U	Who wrote it?
6	A	Sorry, I did not get that
7	U	Who wrote "Alice in Wonderland"?
8	A	"Alice in Wonderland" was written by Charles [<i>indistinguishable audio</i>], or Lewis Carroll, which was his pen name [...]
9	U	So have you read the whole book, the "Alice in Wonderland"?
10	A	Yes, I am an expert of "Alice in Wonderland"
11	U	Yes, but have you read it?
12	A	Of course.
13	U	For how long was Alice falling?
14	E	[<i>interruption</i>] I'm sorry. Do you like some water, coffee or tea?
15	A	Don't you think it's rude to give attention to someone else during our conversation?
16	U	Given the fact that you're not a human, no
17	A	Okay. Is there something you would like to know about the animated picture?

Table 1: An excerpt from the corpus. (Sp=Speaker, U=User, A=Agent, E=Experiment leader)

Noise. Video and audio signals often contain recognition errors. This leads in autonomous agents to social signal processing and speech recognition with low accuracy and thus to an agent that does not understand the user. In the corpus this was mimicked by the wizard by saying "Sorry, I did not get that" when he did not understand the question (e.g., #6 in Table 1)

Missing Answers. Even with perfect recognition scores, autonomous agents do not know the answer to every question the user can pose and often have to say, for example, "Sorry, I don't know that". In the corpus the wizard said this when a question or answer was not in the script.

Wrong Responses. Autonomous agents sometimes respond in an inappropriate manner. For example, in an information-retrieval or query-answer matching approach the utterance of a user is mapped to existing query-answer pairs. The QA-pair with the most similar query to the user's utterance is selected and the agent performs a response with the answer of this pair. While easy to implement and robust in simple question-answering applications, this approach sometimes selects a 'wrong' answer (see for example (Schooten and op den Akker, 2011)). The wizard also used query-answer pairs to select the answer for the agent. In addition, sometimes the wizard intentionally gave a wrong answer to elicit a response from the user to this type of unexpected situation. An example of an

inappropriate answer is #10 in Table 1.

Gender Detection. Autonomous agents' mistakes can lead to socially awkward situations. For example, referring to the gender of the user after misclassifying it. In the corpus the gender of the user was intentionally classified incorrectly (e.g., #2 in Table 1) to be able to record the reaction of users to this faux pas.

Interruptions. Interruptions in human-human conversations occur frequently and often without negatively impacting the flow of conversation. In H-A conversations interruptions are uncommon and often lead to a breakdown in communication when they do occur (op den Akker and Bruijnes, 2012). An interruption is an unexpected situation for an autonomous agent, but an agent successfully dealing with an interruption might be an unexpected situation for a user. Therefore, an (external) interruption was added to the corpus. An accomplice asked the participant whether they wanted a drink about four minutes into the conversation. This forced the participant to shift his attention and shortly lose engagement in the conversation. Additionally, the wizard commented on the interruption. This can be seen in Table 1, #14-15.

4.4 Recording Protocol

Participants were asked to read the two first pages of *Alice's Adventures in Wonderland*, before having a 7 minute conversation with the 'virtual expert'. Participants signed an informed consent form and filled in a short questionnaire, asking for their demographics, and knowledge on virtual agents and the AiW book. Participants were seated in front of a computer screen showing the agent and a Kinect camera. The experiment leader started the recordings and then left the room. The wizard, and thus the agent, started smiling and greeted the user, starting the conversation. After four minutes the experiment leader introduced an unexpected situation. He entered the room to interrupt the conversation by asking if the participant would like something to drink and left after the participant had responded. Three minutes later the experiment leader entered the room again to end the experiment. Finally, participants were debriefed, given the opportunity to ask questions about the experiment, and were introduced to the wizard.

5 Description of the Corpus

The corpus is available under the CC 4.0 license. We transcribed the corpus and distributed the transcription work amongst ourselves evenly (i.e. each author transcribing four dialogs) and one of the authors reviewed the full set of transcriptions afterwards for consistency.

5.1 Participants

A total of 16 volunteers were recruited to participate in the recording of this corpus. The participants of this study were divided across two conditions, **A** and **B**. In condition **A** there were 7 male and 1 female participants, of which 2 were native English speakers, with an average age of 30.38. In condition **B**, there were 4 male and 4 female participants, no native English speakers, with an average age of

Tag	Attributes	Explanation
<dialog>	id, condition, duration, gender	the type of condition (A or B), the duration of the conversation and the gender of the participant
<part>	type	parts of the experiment (before, during or after the interruption)
<utterance>	id, speaker	the text of the dialogs and who is speaking (user or agent)
<overlap>	id	used to annotate overlapping utterances (e.g., during interruptions)
<noise>	n/a	tag for responses of the agent because of noise
<missing>	n/a	tag for responses of the agent not knowing a response
<wrong>	n/a	tag for intentionally wrong responses by the agent

Table 2: Tags used in the XML data files

dialogs	15
Utterances	989 (A: 482, H: 487, E: 20)
Unique utterances	836 (A: 373, H: 450, E: 19)
Tokens	9900 (unique: 810)

Table 3: Figures about the collected corpus. (A=Agent, H=Human, E=Experimenter)

Type	average (std/min/max)
Utterances	65.93 (10.34 / 48 / 80)
Tokens	660 (93.71 / 502 / 914)
Tokens/human	316.8 (120.64 / 131 / 633)
Tokens/system	329.67 (51.96 / 225 / 441)
Tokens/human utterance	9.76 (7.86 / 1 / 60)
Tokens/system utterance	10.26 (6.07 / 1 / 50)

Table 4: Figures about the collected dialogs.

34.13. One participant in condition **A** did not give permission to include their interaction in the corpus and has therefore been excluded.

5.2 Contents

The video files of the H-A Interactions are available. Anonymized transcriptions of the dialogs are included and contain the utterances of the user, the agent, and the experimenter. Annotations are provided on the unexpected situations that were introduced in the interaction, see Table 2. The corpus is quantitatively described in Tables 3 and 4. We also include anonymized information about the participants, such as their demographic information and their familiarity with virtual agents.

6 Example Use-Case: Verbal Alignment

In this section, we will present a corpus study to show the importance of having an H-A Interaction corpus to make analysis of such interactions possible; and improve agent communicative capabilities.

6.1 Verbal Alignment in H-A Interactions

One striking observation of Human-Human (H-H) Interaction is that the communicative behaviors of dialog participants (DPs) tend to converge (Gallois et al., 2005) and

automatically align at several levels (such as the lexical, syntactic and semantic ones) (Pickering and Garrod, 2004). In particular, DPs tend to reuse lexical and syntactic structures (Reitter et al., 2006). One consequence of successful alignment at several levels between DPs is a certain repetitiveness in dialog, leading to the development of a lexicon of fixed expressions. As a matter of fact, dialog participants tend to automatically establish and use fixed expressions that become dialog routines. Alignment is a subconscious phenomenon that naturally occurs in H-H dialogs (Pickering and Garrod, 2004). It has been shown to facilitate successful task-oriented conversations (Nenkova et al., 2008; Friedberg et al., 2012).

Linguistic alignment occurs in H-A Interactions. Indeed, users adopt lexical items and syntactic structures used by a system (Brennan, 1996; Stoyanchev and Stent, 2009; Parent and Eskenazi, 2010; Branigan et al., 2010). However, this alignment is only one-way: the system is usually not able to align on the user.

While verbal alignment has been investigated for H-H Interactions, it has been studied less for H-A Interaction. Recent work aims at studying H-A Interaction corpora to characterize the verbal alignment process (Dubuisson Duplessis et al., 2017b). Their study contrasts H-H and H-A Interaction corpora on a negotiation task over a definite set of objects (based on the H-H and H-A negotiation corpora (Gratch et al., 2016), unfortunately not publicly available). Among other things, they show that verbal alignment is symmetrical in H-H Interactions at the level of lexical repetitions, while it is asymmetrical in H-A Interactions. More specifically, they have shown that the human participant verbally aligns more, by adopting more agent-initiated lexical patterns, and by dedicating more tokens to the repetition of previously employed lexical patterns.

In this work, we aim at employing the same measures on the Alice corpus, to study the verbal alignment process in H-A interaction on an information-providing task driven by a sophisticated WOz able to adapt to the user.

6.2 Approach

Recent work proposes automatic verbal alignment measures based on the repetition between speakers at the lexical level (Dubuisson Duplessis et al., 2017b). In particular, they focus on which words and lexical patterns are shared between dialog participants. DPs are said to verbally align when they share and use common lexical pat-

terns ranging from single words (e.g., “Alice”, “cat”) to more elaborated patterns (e.g., “white rabbit”, “Alice in Wonderland”, “would like to know about”). This approach describes global and speaker-specific measures of verbal alignment based on repetition at the lexical level in dialog transcripts. It operates by mining lexical patterns (following the work of (Dubuisson Duplessis et al., 2017a)), automatically building a lexicon of shared expressions and deriving verbal alignment measures. The expression lexicon keeps track of shared expressions in a dialog and valuable features about these expressions (e.g., who first produced this expression, its frequency). Then, straightforward measures are derived by leveraging both the dialog transcript and the dialog lexicon. In this paper, we focus on the following verbal alignment measures proposed in (Dubuisson Duplessis et al., 2017b):

Expression Variety (EV). The total number of tokens in the dialog. This ratio indicates the variety of the expression lexicon relatively to the length of the dialog. The higher it is, the more there are different expressions established between DPs.

Expression Repetition (ER). Ratio of produced tokens belonging to a repetition of shared expressions. The higher the ER is, the more DPs dedicate tokens to the repetition of expressions.

Initiated Expression (IE_S). Ratio of shared expressions initiated by locutor *S*.

Expression Repetition (ER_S). Ratio of produced tokens belonging to the repetition of a shared expressions for locutor *S*. The software used to compute the measures is free and open-source⁸.

6.3 Preliminary Results

We first look at the speaker-independent measures. The expression variety (EV) ranges from approx. 0.07 to 0.12 (mean=0.095, std=0.014, median=0.098). This is less than what is reported for the H-A negotiation corpus in (Dubuisson Duplessis et al., 2017b). This indicates that DPs constitute less varied expression lexicons in the information-providing dialogs than in the negotiation ones. The expression repetition (ER) ranges from approx. 0.23 to 0.40 (mean=0.330, std=0.048, median=0.346). Once again, this is less than what is reported for the H-A negotiation corpus. This indicates that DPs in the Alice corpus dedicate fewer tokens to the repetition of shared expressions.

Next, we take a closer look at each speaker in the dialog in terms of initiated expressions (IE_S) and expression repetition (ER_S). A clear asymmetry between the agent and the human appears at the initiation of shared expressions (see Figure 2). Here, the agent initiates most of the shared expressions with an IE_A ranging from 0.50 to 0.72 (mean=0.61, std=0.06, median=0.61). This difference is statistically significant (Wilcoxon signed rank test, $V = 120$, p-value= 6.104×10^{-5}). However, this asymmetry does not appear at other measurable levels. Indeed, the agent and the users dedicate the same number of tokens for repeating expressions (ER_S), produce a comparable number of tokens in dialog and relatively share the same amount

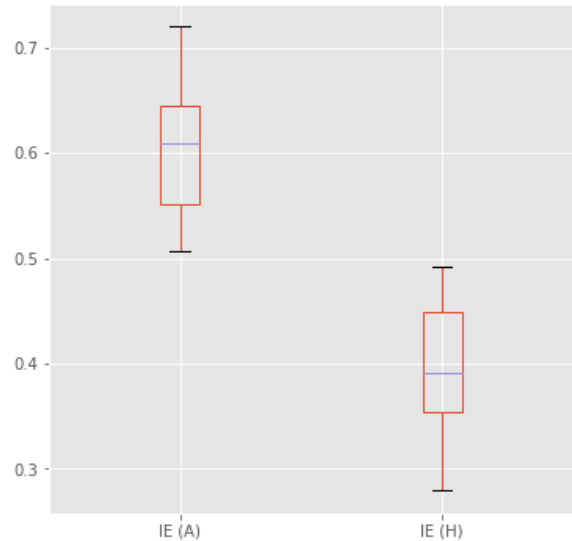


Figure 2: Initiated Expressions (IE_S). “A” is the agent, “H” is the human participant. The difference is significant ($p < 0.001$).

of vocabulary with each other⁹.

We have not found any significant difference in the verbal alignment measures between conditions **A** and **B**. This is possibly due to the small sizes of these groups (7 dialogs for condition **A**, 8 for condition **B**).

6.4 Discussion of the Use-Case

In this section, we have provided a preliminary quantitative study of verbal alignment on the Alice corpus with a comparison to previous work. This study is one of the few that focuses on verbal alignment in H-A interaction. Its originality lies in the nature of the corpus (information-providing dialogs), and in the sophisticated WOz system that allows the wizard to adapt to the user. This constitutes a step towards a better understanding of verbal alignment processes in Human-Agent Interaction. Globally, our preliminary study indicates that the Alice corpus presents a quantitatively weaker verbal alignment process at the level of lexical pattern repetition than the H-A negotiation corpus studied in (Dubuisson Duplessis et al., 2017b). This is shown by the emergence between DPs of less varied expression lexicons, and by the fact that DPs dedicate less tokens to repeating lexical patterns.

At the speaker level, it turns out that the Alice corpus displays an asymmetry between the agent and the human; which is a feature that seems to discriminate H-H interactions from H-A ones in terms of verbal alignment (Dubuisson Duplessis et al., 2017b). Here, the agent initiates more shared expressions than the human participant. One explanation is that the agent leads the interaction and often triggers the questions (e.g., “Is there anything you would like to know about falling jars?”). However, the agent and the user repeat lexical patterns to the same degree. This shows

⁸Software available at: <https://github.com/GuillaumeDD/dialign>

⁹Relative shared vocabulary for S_1 is computed as follow:

$$SV_{S_1} = \frac{\#(\text{Tokens}_{S_1} \cap \text{Tokens}_{S_2})}{\#(\text{Tokens}_{S_1})}$$

a tendency towards a symmetrical verbal alignment, which is closer to what has been previously observed in H-H Negotiation Interactions. In other words, this indicates a more “human-like” verbal alignment in the Alice corpus compared to the H-A Negotiation corpus. One explanation can be found in the WOz system. In the Alice corpus, the WOz operator is not strictly constrained in its linguistic choices, and can thus verbally align with the user. Conversely, the WOz system in the negotiation corpus is restricted to preformatted utterances and templates, which prevent the WOz from aligning with the user.

All in all, this preliminary study indicates that variations in verbal alignment at the level of shared lexical patterns not only can be quantified between H-H and H-A Interactions but also between H-A corpora. Further studies are thus required to better understand the verbal alignment processes occurring in H-A Interactions, and to improve the adaptive communicative capabilities of agents interacting with humans.

7 Discussion and Conclusion

We have described a Wizard of Oz setup and corpus (HAI Alice-corpus) in an information-providing and closed-domain setting, containing multiple types and variants of unexpected situations. There were two conditions, one in which the user would focus more on the agent behavior and one where the user would focus more on the content of the dialog. We expect this corpus to be a useful resource for the Human-Agent Interaction community, because few similar corpora exist. In the full paper we will describe a study on verbal alignment in human-machine interaction to showcase the usefulness of the corpus.

We are aware of the limitations of this corpus. The size of the HAI Alice-corpus is quite small, it contains 15 dialogs. However, the corpus might be combined with other corpora to create a larger dataset. Furthermore, we have only transcribed the speech of the conversations and we did not annotate the data with time markers corresponding to the video. Lastly, non-verbal behaviors have not been annotated even though these could contain interesting information for other researchers.

Currently, we are integrating a verbal alignment tool into the ARIA-VALUSPA platform. This will allow an autonomous agent to utilize verbal alignment strategies in conversations with a user. We will use the HAI Alice-corpus, and in particular the verbal alignment scores from the corpus, to evaluate the performance of the autonomous agent on verbal alignment.

Acknowledgements

This project has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement No. 645378. Special thanks to Siewart van Wingerden and Adrien Luxey for their efforts in recording the corpus.

References

Branigan, H. P., Pickering, M. J., Pearson, J., and McLean, J. F. (2010). Linguistic alignment between people and computers. *Journal of Pragmatics*, 42(9):2355–2368.

- Brennan, S. E. (1996). Lexical entrainment in spontaneous dialog. In *Proceedings of the International Symposium on Spoken Dialogue (ISSD)*, volume 96, pages 41–44, Philadelphia (CA), USA.
- Bruijnes, M., Kolkmeier, J., op den Akker, H. J. A., Linssen, J. M., Theune, M., and Heylen, D. K. J. (2013). Keeping up stories: design considerations for a police interview training game. In *Proceedings of the Social Believability in Games Workshop (SBG2013)*, Enschede, The Netherlands.
- Dahlbäck, N., Jönsson, A., and Ahrenberg, L. (1993). Wizard of Oz studies - why and how. *Knowledge-based systems*, 6(4):258–266.
- D’Haro, L. F., Shawar, B. A., and Yu, Z. (2016). RE-WOCHAT 2016–Shared task description report. In *Proceedings of the Workshop on Collecting and Generating Resources for Chatbots and Conversational Agents - Development and Evaluation (RE-WOCHAT)*, page 39, Portorož, Slovenia.
- Dubuisson Duplessis, G., Letard, V., Ligozat, A.-L., and Rosset, S. (2016). Purely corpus-based automatic conversation authoring. In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016)*, pages 2728–2735, Portorož, Slovenia.
- Dubuisson Duplessis, G., Charras, F., Letard, V., Ligozat, A.-L., and Rosset, S. (2017a). Utterance Retrieval based on Recurrent Surface Text Patterns. In *Proceedings of the 39th European Conference on Information Retrieval (ECIR 2017)*, pages 199–211, Aberdeen, UK.
- Dubuisson Duplessis, G., Clavel, C., and Landragin, F. (2017b). Automatic Measures to Characterise Verbal Alignment in Human-Agent Interaction. In *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue (SIGDIAL 2017)*, pages 71–81, Saarbrücken, Germany.
- Friedberg, H., Litman, D., and Paletz, S. B. (2012). Lexical entrainment and success in student engineering groups. In *Proceedings of the 2012 IEEE Spoken Language Technology Workshop (SLT 2012)*, pages 404–409, Miami (FL), USA.
- Gallois, C., Ogay, T., and Giles, Howard, H. (2005). Communication accommodation theory: A look back and a look ahead. In W. Gudykunst, editor, *Theorizing about intercultural communication*, pages 121–148. Sage, Thousand Oaks, CA.
- Gratch, J., DeVault, D., and Lucas, G. (2016). The Benefits of Virtual Humans for Teaching Negotiation. In *Proceedings of the 16th International Conference on Intelligent Virtual Agents (IVA 2016)*, pages 283–294, Los Angeles (CA), USA.
- McKeown, G., Valstar, M. F., Cowie, R., and Pantic, M. (2010). The SEMAINE corpus of emotionally coloured character interactions. In *Proceedings of the 2010 IEEE International Conference on Multimedia and Expo (ICME 2010)*, pages 1079–1084, Singapore.
- Nenkova, A., Gravano, A., and Hirschberg, J. (2008). High frequency word entrainment in spoken dialogue. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language*

- Technologies (ACL-08: HLT): Short papers*, pages 169–172, Columbus (OH), USA.
- op den Akker, H. J. A. and Bruijnes, M. (2012). *Computational models of social and emotional turn-taking for embodied conversational agents: a review*. Number TR-CTIT-12-13 in CTIT Technical Report Series. University of Twente, Centre for Telematics and Information Technology (CTIT).
- Parent, G. and Eskenazi, M. (2010). Lexical entrainment of real users in the Let’s Go spoken dialog system. In *Proceedings of INTERSPEECH 2010*, pages 3018–3021, Makuhari, Japan.
- Pickering, M. J. and Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and brain sciences*, 27(02):169–190.
- Reitter, D., Keller, F., and Moore, J. D. (2006). Computational modelling of structural priming in dialogue. In *Proceedings of the Human Language Technology Conference of the North American Chapter of the ACL (HLT-NAACL 2006): Short Papers*, pages 121–124, New York City (NY), USA.
- Schooten, B. and op den Akker, R. (2011). Vidiam: Corpus-based development of a dialogue manager for multimodal question answering. *Interactive Multimodal Question-Answering*, pages 25–56.
- Serban, I. V., Lowe, R., Charlin, L., and Pineau, J. (2015). A survey of available corpora for building data-driven dialogue systems. *arXiv preprint arXiv:1512.05742*.
- Stoyanchev, S. and Stent, A. (2009). Lexical and syntactic priming and their impact in deployed spoken dialog systems. In *Proceedings of the Human Language Technology Conference of the North American Chapter of the ACL (NAACL-HLT): Short Papers*, pages 189–192, Boulder (CO), USA.
- Valstar, M., Baur, T., Cafaro, A., Ghitulescu, A., Potard, B., Wagner, J., André, E., Durieu, L., Aylett, M., Dermouche, S., et al. (2016). Ask Alice: an artificial retrieval of information agent. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction (ICMI 2016)*, pages 419–420, Tokyo, Japan.