
Investigating syntactic alignment in spoken natural language human-computer communication

Benjamin R. Cowan

School of Computer Science
The University of Birmingham
Birmingham, UK
B15 2TT
b.r.cowan@cs.bham.ac.uk

Russell Beale

School of Computer Science
The University of Birmingham
Birmingham, UK
B15 2TT
R.Beale@cs.bham.ac.uk

Holly P. Branigan

Psychology Department
The University of Edinburgh
Edinburgh, UK
EH8 9JL
holly.branigan@ed.ac.uk

Copyright is held by the author/owner(s).
CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.
ACM 978-1-4503-0268-5/11/05.

Abstract

This paper describes planned experiment-based research observing the existence of syntactic alignment in natural language computer interactions. This research will achieve this through using a computer-human version of the confederate communication task commonly used in psycholinguistic research observing syntactic alignment in human-human dialogue. The motivations of the work lie in observing the existence of syntactic alignment in human-computer dyads and how the naturalness of interaction affects the appearance of such a linguistic phenomenon. The work will also aim to identify how such a linguistic effect links to users' satisfaction and quality judgments of interaction.

Keywords

Syntactic Alignment, Natural Language Dialogue Systems, Psycholinguistics, HCI, User Satisfaction, Metric, Experiment-based research

ACM Classification Keywords

H5.2 User Interfaces: Natural Language, Evaluation/Methodology

General Terms

Human Factors, Experimentation

Introduction

Humans are intrinsically social beings and language is at the core of most social activity. We use language to describe ideas, thoughts and feelings to others on a frequent basis. Not only do we direct speech at others, others' speech directed to us also affects how we create our utterances reflecting a linguistic concept termed *alignment*. This alignment is a dynamic process by which we converge on shared mental representations within dialogue affecting the words and grammatical structures we use when conversing with a conversational partner [2]. Alignment during dialogue can include lexical or syntactic alignment, and such effects have been consistently demonstrated in research with human interlocutors (conversational partners) [12,13]. Alignment is argued to be integral in successful communication between conversational partners [8].

Recent interest in this area has shifted from observing this alignment effect within human conversational dyads towards the influence and existence of alignment during human-computer interactions [2,3,4,11]. Such work has shown that linguistically users tend to align in computer interaction and that this effect can also be more pronounced in human-computer linguistic interactions. Recent research [3] investigated the impact of verb repetition within the interlocutors' descriptions on alignment between human-human and human-computer dyads. Participants were asked to play a picture matching game with a conversational partner (a computer instead of a human confederate as used in the confederate scripting paradigm in human-human alignment research [1]). Both partners took turns to describe images using text for the other to match with their pictures on screen. In one condition the

participant was led to believe that they were interacting with a computer whereas in the other condition they were led to believe they were interacting with a human partner (they were in fact interacting with a computer in both). It was found that beliefs about the interlocutor influenced the amount of alignment present in the participants' sentences. Participants aligned more in the computer-human interaction rather than human-human interaction when verbs used by participants' were the same to those previously used by the confederate in the computer condition.

Additionally research in this field has also suggested that expectations towards system abilities and quality influences users' lexical alignment. When playing a collaborative computer-based picture matching game with a computer partner users were made to believe that they were interacting with either a "basic" or "advanced" computer. Interestingly although users aligned more towards the computers lexical representation in both conditions, this alignment was significantly larger when interacting with the "basic" computer [11]. This phenomenon of heightened alignment seen in both [3] and [11] is likely due to the speakers' beliefs of the system function and from this their ideas on the behaviour needed to enhance communication effectiveness, an example of mediated alignment (alignment mediated by beliefs about the audience, here analogous to the psycholinguistic concept of audience design in dialogue). Alignment has also been considered to occur unconsciously (unmediated by beliefs about the audience), due to a priming effect on mental linguistic representations in language processing. Specifically in syntactic alignment, activation of grammatical nodes (representations) in the lemma stratum from utterance processing in production or comprehension is thought

to heighten activation of these nodes in subsequent language use, leading to a higher likelihood of using the same form in speaker utterances [1,6]. Additionally social influences such as politeness and the desire to create a social bond [5] also influence alignment. As computers have been demonstrated to be social actors [14] such alignment due to social concepts may also be apparent in HCI.

As can be seen from the research reviewed alignment does occur in HCI, yet the impact of natural dialogue between computers and humans and the effect that this has on alignment processes has not been explored. Much of the research on alignment in HCI has been based on textual input rather than natural dialogue interaction with a speech interface. Studying syntactic alignment in a spoken interaction is important due to the core differences in communication through text and spoken dialogue in terms of dynamism, spontaneity and social information present in each. Additionally much research has concentrated on lexical alignment (the alignment of word choices between interlocutors, which may reflect alignment of meaning representations) rather than syntactic alignment (reflection of grammatical representations within dialogue) in human-computer dialogue interactions. This work aims to research syntactic alignment in natural dialogue interactions, something that to the authors' knowledge has not been studied. An interaction with a truly advanced computer with natural language capabilities may not only potentially lead to a development of bond but also a shift in the user model of computer function away from perceived limitations of the computer system towards more natural conversation-based levels. In more natural interaction there may therefore be a tipping point whereby alignment switches from being essentially

mediated by beliefs about the audience to unmediated as the blurring between computer and human in representations occurs.

Additionally as capability grows for computers to communicate naturally with humans, research on the alignment of dialogue and how this relates to interaction quality could be useful in the assessment of interactions with such natural systems. Alignment may index not only the communicative success but also the rapport between the agent and the human. However this relationship may not be a straightforward linear one. For example, alignment up to a certain degree may be a positive marker, reflecting an increasing social bond. Yet beyond this level, overly strong alignment may instead reflect users' overcompensation for perceived or actual deficiencies in system function. It may be that the levels of alignment naturally found in human interaction reflect the levels that yield optimal perceptions of naturalness and satisfaction. To use a relevant context, linguistic alignment between a virtual agent using natural language and the human user may be an indication not only of the communicative success but the rapport between the agent and the human. Strong alignment (i.e. that above levels seen in human to human dialogue) however might reflect overcompensation in the human users behaviour be it due to their beliefs about the system or actual poor system function, so that they do not feel the interaction to be natural and satisfying.

Research Aims and Impact

From the planned experiment described below we aim to gather insight into the impact of natural dialogue interaction (rather than written utterances) on syntactic (rather than lexical) alignment in HCI. It also aims to investigate how this alignment may relate to users'

satisfaction with natural dialogue systems interactions. This will shed light on how important alignment is for subjective interaction quality whilst also exploring whether syntactic alignment could potentially be used as a behavioural metric of satisfaction or achieved naturalness of such systems. Such a measure will be useful in its ability to infer quality of linguistic interaction compared to the use of more subjective satisfaction questionnaires currently used in HCI research. The research is also forward-looking in the fact that existing natural language interfaces are not sufficiently complex or adaptive to offer the subtlety and flexibility required in natural dialogue. This work will therefore inform how systems should ideally behave once the technology is sufficiently developed. From these findings the research would look to develop a deeper understanding of syntactic alignment in dialogue HCI and how this could be considered and observed in a fully operational natural language dialogue system interaction.

Proposed Programme

This work is investigating the effect of alignment so that, when systems' natural linguistic interaction capabilities are more developed we will actually know what we want the system to achieve. The proposed research being undertaken will aim to use a wizard of oz experiment methodology to observe user syntactic alignment in natural dialogue between a computer and a human. A wizard of oz methodology involves the simulation of a system interaction rather than interacting with a final system. Such methods tend to be used to simulate future technologies or functionality of technologies not currently developed to high enough standards to simulate the desired interaction. At present existing systems are not sufficiently flexible or

intelligent enough to offer the full subtleties required to observe syntactic alignment for the aims above and hence we need to use a wizard of oz methodology to effectively study this. The planned experiment will use a between participants design. The experiment design to be used is similar to that used in existing human-human interaction research in dialogue. The planned experiment holds three conditions used to explore this phenomenon. The first condition involves participants interacting with a confederate to observe syntactic alignment in human-human interaction. This will act as a control condition to compare the effects seen in human-human dyads to computer-human dyads. The second condition will see users interact with a limited computer interlocutor. The speech created will not include features regular to human speech such as intonation, accent and varying speech rate likely to influence the naturalness of dialogue. The third condition will involve users interacting with a more linguistically natural computer imitated by using recordings made by a human confederate.

The task used will be a confederate-based communication task commonly used in syntactic alignment in dialogue research whereby two players (the naïve participant and a confederate) take turns in describing and matching different picture cards in a collaborative communication game [1,6]. The confederate (either computer or human depending on the conditions) will produce predefined utterances (prime utterances) varying in grammatical structure and the subsequent utterances by the participant (the target utterance) will be observed to identify whether alignment has occurred.

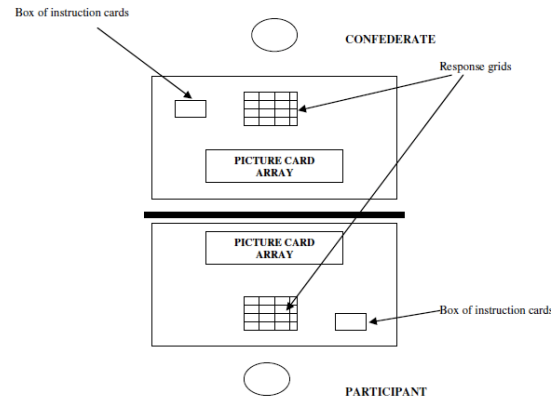


Figure 1: Referential Communication Task Set Up (Control condition)

Each picture card being described in the experiment varies in colour (blue, orange, yellow and purple), shape (square, star, triangle and circle) and pattern (stripey, dotted, wavy and chequered) combination. The confederate produces specified scripted noun phrase orders that vary in grammatical structure (e.g. “purple stripey square”, “stripey purple square” or “the square that’s purple and stripey”). The order of the descriptions will be controlled to ensure that the nouns used in the participant’s utterance will not be the same as that used in the priming phrase from the confederate. This is to control for potential alignment due to lexical similarity. If syntactic alignment occurs then these will be reflected in the utterances of the conversational partner when they describe their cards. These interactions will be recorded and coded in terms of whether the participant utterance (target) is the same or different when compared to the confederate (prime) utterance in grammatical structure.

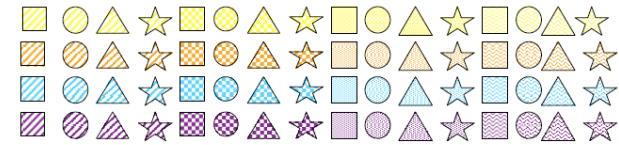


Figure 2: Example of the full picture card set used in the experiment

Satisfaction with the interaction will also be measured and analysed in the research and forms a crucial part of the research aims (as described previously). The research will measure this after interaction using a satisfaction measure created specifically for this interaction using themes of items from the MINERVA [10] and SUMI [9] satisfaction self-report scales. The SUMI [9] is a 50 item 3-point Likert scale questionnaire used to measure software usability. It is constructed of 5 factors referring to users ratings of efficiency, affect, helpfulness, control and learnability of the software. The MINERVA usability questionnaire was originally designed to measure usability of dialogue systems making the item content highly relevant to this research. It uses a 7-point Likert scale with which users note their level of agreement with each item. It has been successfully used as a usability measure in dialogue system interactions [7]. The items focus around the factors of cognitive effort and stress, fluency of interaction, transparency, interface quality and engagement/enjoyment within the interaction. The data from this will be used to investigate how alignment relates to user satisfaction in the experiment conditions and how naturalness of the interaction affects satisfaction.

An interview at the end of the experiment will also be used to gather more qualitative data about their

opinions on the interactions and to gather ideas for further research.

Acknowledgements

This research is funded by the University of Birmingham's Ramsay Fund. The authors would also like to thank the anonymous reviewers for their comments and suggestions.

References

1. Branigan, H.P., Pickering, M.J., and Cleland, A. Syntactic co-ordination in dialogue. *Cognition* 75, (2000), B 13-25.
2. Branigan, H.P., Pickering, M.J., Pearson, J., and McLean, J.F. Linguistic alignment between people and computers. *Journal of Pragmatics* 42, 9 (2010), 2355-2368.
3. Branigan, H.P., Pickering, M.J., Pearson, J.M., McLean, J.F., and Nass, C. Syntactic alignment between computers and people: the role of belief about mental states. *Proceedings of the Twenty-fifth Annual Conference of the Cognitive Science Society*, Erlbaum, Mahwah (2003), 186-191.
4. Brennan, S.E. Lexical entrainment in spontaneous dialog. *Proceedings of the International Symposium on Spoken Dialogue*, (1996), 41-44.
5. Chartrand, T. and Bargh, J. The chameleon effect: The perception-behaviour link and social interaction. *Journal of Personality and Social Psychology* 76, (1999), 893-910.
6. Cleland, A. and Pickering, M.J. The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language* 49, 2 (2003), 214-230.
7. Davidson, N., McInnes, F., and Jack, M.A. Usability of dialogue design strategies for automated surname capture. *Speech Communication* 43, (2004), 55-70.
8. Garrod, S. and Pickering, M. Joint Action, Interactive Alignment, and Dialogue. *Topics in Cognitive Science* 1, (2009), 292-304.
9. Kirakowski, J. and Corbett, M. SUMI: the software usability measurement inventory. *British Journal of Educational Technology* 24, (1993), 210-212.
10. Love, S., Dutton, R.T., Foster, J.C., Jack, M.A., and Stentiford, F.W.M. Identifying salient usability attributes for automated telephone services. *Proceedings of the 3rd International Conference on Spoken Language Processing*, ISCA (1994), 1307-1310.
11. Pearson, J., Hu, J., Branigan, H.P., Pickering, M.J., and Nass, C. Adaptive language behavior in HCI. *Proceedings of the SIGCHI conference on Human Factors in computing systems - CHI '06*, (2006), 1177.
12. Pickering, M.J. and Branigan, H.P. The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language* 39, (1998), 633-651.
13. Pickering, M.J. and Ferreira, V.S. Structural Priming: A Critical Review. *Psychological Bulletin* 134, 3 (2008), 427-459.
14. Reeves, B. and Nass, C. *The Media Equation: How People Treat Computers, Television, and New Media like Real People and Places*. Cambridge University Press, 1998.