



Voice anthropomorphism, interlocutor modelling and alignment effects on syntactic choices in human–computer dialogue [☆]



Benjamin R. Cowan ^{a,c,*}, Holly P. Branigan ^b, Mateo Obregón ^b, Enas Bugis ^a, Russell Beale ^a

^a HCI Centre, School of Computer Science, University of Birmingham, Edgbaston Campus, Birmingham B15 2TT, England

^b Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, Scotland

^c School of Information and Library Studies, University College Dublin, Belfield, Dublin 4, Ireland

ARTICLE INFO

Article history:

Received 26 June 2014

Received in revised form

1 May 2015

Accepted 28 May 2015

Communicated by Pelachaud Catherine

Available online 7 June 2015

Keywords:

Human–computer dialogue

Syntactic alignment

Speech interaction

User behaviour

Psycholinguistics

Interlocutor modelling

ABSTRACT

The growth of speech interfaces and speech interaction with computer partners has made it increasingly important to understand the factors that determine users' language choices in human–computer dialogue. We report two controlled experiments that used a picture-naming-matching task to investigate whether users in human–computer speech-based interactions tend to use the same grammatical structures as their conversational partners, and whether such *syntactic alignment* can impact strong default grammatical preferences. We additionally investigate whether beliefs about system capabilities that are based on partner identity (i.e. human or computer) and speech interface design cues (here, voice anthropomorphism) affect the magnitude of syntactic alignment in such interactions. We demonstrate syntactic alignment for both dative structures (e.g., *give the waitress the apple vs. give the apple to the waitress*), where there is no strong default preference for one or other structure (Experiment 1), and noun phrase structures (e.g., *a purple circle vs. a circle that is purple*), where there is a strong default preference for one structure (Experiment 2). The tendency to align syntactically was unaffected by partner identity (human vs. computer) or voice anthropomorphism. These findings have both practical and theoretical implications for HCI by demonstrating the potential for spoken dialogue system behaviour to influence users' syntactic choices in interaction. As well as verifying natural corpora findings, this work also highlights that priming and cognitive mechanisms that are unmediated by beliefs about partner identity could be important in understanding why people align syntactically in human–computer dialogue.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Recent innovations in consumer electronics have led to a rapid increase in the frequency of spoken dialogue interactions between people and artificial systems, where users use natural speech to command devices or to query systems, and devices and systems in turn use natural speech to respond. Speech and human–computer dialogue interactions are now common in mainstream technology products; 87% of iPhone 4S users have reported using Siri at least once a month (Barrett and Jiang, 2012) and many other services such as Google Now, in-car systems and Smart TVs are using speech as an interaction modality. The future growth of human–robot interaction as well as the use of embodied conversational agents highlights that natural dialogue interactions between computers and humans are likely to become more prominent. With this in mind, recent calls have

been made for HCI and speech-based researchers to combine efforts to understand what governs our interactions with speech technology to design more effective speech interface interactions (Aylett et al., 2014).

Although a vast body of knowledge has been accumulated about the factors that govern spoken dialogue behaviours between two people (*human–human dialogues*; *HHD*), comparatively little is known about spoken dialogues between people and artificial systems (*human–computer dialogues*; *HCD*). In this paper, we focus on one particular factor that has been hypothesised to strongly influence speakers' behaviour in HHD: linguistic *alignment*, or the tendency for conversational partners to converge on common language choices.

We investigate whether users show alignment of grammatical structure (*syntactic alignment*) in speech-based HCDs under controlled experimental conditions using a game in which a participant and their partner alternately describe and match pictures. The game includes images that can be described by two different grammatical structures. In the game, the participant's partner (a 'confederate') uses specific grammatical structures (*primes*) to describe their images and we observe whether participants then tend to use the same structure

[☆]This paper has been recommended for acceptance by Henrik Iskov Christensen.

* Corresponding author at: School of Information and Library Studies, University College Dublin, Belfield, Dublin 4, Ireland. Tel.: +353 17167660.

E-mail address: benjamin.cowan@ucd.ie (B.R. Cowan).

rather than the alternative structure in their subsequent descriptions. The game therefore gives participants an opportunity to make choices around the structure they use to describe the images whilst allowing experimenters to observe the effect of the prime on their syntactic choice. Although previous more naturalistic research has suggested that users do align syntactically in speech-based HCD (Stoyanchev and Stent, 2009), such work does not control potential confounds that may affect the magnitude of alignment (e.g. effects associated with repetition of particular words, recency effects, natural frequency of structures and speech recognition errors). Using a controlled experimental paradigm such as the picture-description-matching game allows a precise focus on the causal impact of particular variables while controlling such confounds in the game materials. Indeed such studies are important in validating more naturalistic corpora work (Gilquin and Gries, 2009). The current study also expands previous laboratory-based alignment research on text (Branigan et al., 2003) to speech-based dialogue interactions, reflecting the increased prominence of speech as an interaction modality in popular devices. Previous research has highlighted higher sharing of syntax with the computer in speech-based compared to text-based interactions (Le Bigot et al., 2007). Text-based studies of alignment may therefore underestimate the magnitude of alignment in spoken HCD, making it important to explore syntactic alignment in a speech-based HCD context.

Importantly, we examine alignment for two types of syntactic alternation that differ in their default structural preferences. Experiment 1 focuses on Double Object (DO; e.g. “The cowboy offering the robber the banana”) and Prepositional Object (PO; e.g. “The cowboy offering the banana to the robber”) structures. These are relatively evenly balanced in terms of default preferences in a non-biasing context¹ when people are describing dative events (i.e., events involving transfer of possession; roughly 60% PO, 40% DO, based on natural language corpora: Gries and Stefanowitsch, 2004; see also Pickering et al., 2002 for similar evidence from experimental studies). Experiment 2 focuses on noun phrase structures, specifically Adjective-Noun (AN; e.g., *the red circle*) and Noun-Relative Clause (RC; e.g., *the circle that's red*) structures. People have been shown to have a strong default preference for using AN structures in a non-biasing context when they are describing relevant items such as coloured and patterned shapes (around 95%, based on available evidence from experimental studies; Branigan et al., in preparation). Studying two syntactic alternations allows us to verify the generalizability of our findings and test whether mechanisms of syntactic alignment in HCI are sufficiently influential to impact strong intrinsic structural preferences. In addition it also allows us to explore whether syntactic alignment might be mediated more strongly by beliefs about the conversational partner, or *interlocutor*, for structure choices in which one alternative is strongly favoured (as in noun phrase structures). Under these circumstances, the choice between a strongly favoured and a strongly disfavoured alternative might be particularly salient, and might therefore be more amenable to strategic decisions based on beliefs about interlocutors' likely understanding or preferences.

The work also adds insight onto the role that partner type (i.e. computer or human) and design choices (such as voice anthropomorphism) have on levels of syntactic alignment. Current findings in HCD suggest that many of our language behaviours are mediated by our perceptions of computers as effective communication partners (Amalberti et al., 1993; Bell and Gustafson, 1999; Brennan, 1998; Le Bigot et al., 2007). In particular, research on alignment of lexical choice in HCD (Bergmann et al., in press; Branigan et al., 2011) suggests that users adapt their lexical choices to accommodate their partner's

perceived limitations as an interlocutor, with greater adaptation to partners perceived as less able. Work on anthropomorphic robotic agents suggests that we see such agents as more intelligent and capable than non-anthropomorphic agents (Kiesler et al., 2008; King and Ohya, 1996). This raises the possibility that anthropomorphic cues in HCD scenarios may lead users to adapt less in these contexts than when interacting with a computer partner with less anthropomorphic cues.

Validating the occurrence of syntactic alignment in speech-based HCD under controlled experimental conditions, and demonstrating that characteristics of the partner affect syntactic alignment, would provide evidence that computer partner utterances as well as design can act as a means of inducing users to use predictable structures that the system can process successfully. Importantly the work also has implications for the understanding of what guides our linguistic choices in HCD. Demonstrating that syntactic alignment is impacted by the anthropomorphism of the partner (and indeed by whether the partner is a computer or human) would show that syntactic alignment, like other language behaviours in HCD, is adaptive and influenced by our perceived limitations of the system as a dialogue partner. In contrast, if we found that syntactic alignment occurs in HCD but is unaffected by partner type, this would tentatively support the notion that cognitive architectures involved in language comprehension and production and the priming of language representations may play a role in syntax choice in HCD (Pickering and Garrod, 2004).

In both experiments, native English speakers played a picture-naming and -matching game with either a human, a computer with an anthropomorphic voice, or a computer with a robotic voice. The computer voices used were shown to yield significant differences in user perceptions of partner ability, with the anthropomorphic voice leading people to see a computer partner as being more advanced, flexible and competent than those hearing the robotic voice (see Section 3). The participants and their partners took turns describing pictures of dative events (Experiment 1) or colored patterned shapes (Experiment 2), and choosing pictures in response to their partner's descriptions (in all experiments the partner was either a human confederate or a computer controlled remotely by a member of the experiment team). Two types of structural alternation were tested across the experiments, each with two alternatives that were primed using the partner's descriptions (Experiment 1: Prepositional Object (PO) or Double Object (DO); Experiment 2: Adjective-Noun (AN) or Noun-Relative clause (RC)) with the experimenter noting the structure that the participants used when producing their own immediately subsequent description. Syntactic alignment is said to occur when participants used the same structure as the prime they were previously exposed to. We found that syntactic alignment occurred in both experiments, yet this effect was not significantly impacted by the partner conditions. This supports the notion that automatic priming of linguistic representations may play a significant role in alignment of syntax choice in HCD.

2. Background

2.1. Alignment in human–human dialogue

A large body of evidence from HHD has shown that conversational partners influence each other's behaviour. In particular, conversational partners show a robust tendency to converge on, or *align*, their non-linguistic and linguistic behaviour, such as posture and gestures (Chartrand and Bargh, 1999; Van Baaren et al., 2009), as well as semantic, lexical, and syntactic choices (Branigan et al., 2000; Brennan and Clark, 1996; Clark and Brennan, 1991; Garrod and Anderson, 1987; Pickering and Branigan, 1998). Alignment of language has been hypothesised to play a causal role in successful communication: by aligning their linguistic representations in production and

¹ In this case, a non-biasing context refers to a context where people have not just been systematically exposed to one or other structure produced by their conversational partner (i.e. a prime).

Download English Version:

<https://daneshyari.com/en/article/401119>

Download Persian Version:

<https://daneshyari.com/article/401119>

[Daneshyari.com](https://daneshyari.com)