Applied Psycholinguistics (2025), **46**, e22, 1–37
doi:10.1017/S0142716425100118

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

Linguistic alignment of redundancy usage in human-human and human-computer interaction

Max S. Dunn[1] and Zhenguang G. Cai[2,3]

[1]Department of English and Communication, The Hong Kong Polytechnic University, Hong Kong, SAR, China, [2]Department of Linguistics and Modern Languages, The Chinese University of Hong Kong, Hong Kong, SAR, China and [3]Brain and Mind Institute, The Chinese University of Hong Kong, Hong Kong, SAR, China
Corresponding author: Max S. Dunn; Email: max-shaw.dunniii@polyu.edu.hk

(Received 9 July 2024; revised 23 June 2025; accepted 10 July 2025)

Abstract
While speakers are theorized to ideally not include unnecessary information (redundancy) in their utterances, in reality, they often do so. One potential reason is that linguistic redundancy facilitates language communication, especially when the addressee (interlocutor) is linguistically less competent (e.g., an artificial system). In three experiments, we examined whether linguistic redundancy may arise as a result of people's tendency to use similar linguistic features as their interlocutor does during communication (i.e., linguistic alignment) and whether redundancy alignment (if any) differs with a human interlocutor versus a computer interlocutor. We also examined whether redundancy alignment is affected by the perceived competency of the interlocutor, participants' abilities in theory of mind (ToM), and if redundancy alignment varied across time during the experiment. Participants carried out a picture matching and naming task with a human or computer interlocutor who either always or never included redundancies in their utterances. Redundancy alignment was found across all experiments, in that speakers produced more redundancies with a redundant interlocutor compared to a non-redundant one. This alignment was also modulated by the perceived competency of the interlocutor, the time course of the interaction, and ToM abilities, suggesting that redundancy usage is affected by both automatic and strategic mechanisms of linguistic alignment.

Keywords: Human-computer interaction; language production; linguistic alignment; redundancy

Introduction

Speakers are theorized to communicate efficiently, avoiding unnecessary redundant linguistic information whenever possible (Grice, 1975). However, in actual language use, speakers sometimes include redundant information in their utterances, for

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

example, referring to a green apple as *the green apple* instead of *the apple* when there is only one apple in the scene (e.g., Engelhardt et al., 2006). This redundancy may occur because such expressions provide additional cues (e.g., color) that facilitate language communication (Deutsch & Pechmann, 1982; Rubio-Fernández, 2021; Saryazdi et al., 2022) or because speakers adopt similar expressions used by their conversation partner (i.e., their interlocutor), such as using color in naming (Branigan et al., 2004). These linguistic redundancies can be particularly important when interacting with someone perceived to have limited language competence, such as a child or a computer. Our study explores whether speakers align with their interlocutor's redundancy patterns and whether such alignment is influenced by beliefs about the interlocutor when interacting with a human or a computer.

Linguistic redundancy

When using language, even simple descriptions of objects or events can be formulated in many different ways. For example, to refer to an apple, a speaker could use phrases such as *the apple*, *the green apple*, *the apple on the towel*, etc. It is typically assumed that speakers follow the Maxim of Quantity (Grice, 1975), ensuring that their utterances provide sufficient information for successful communication without including redundant information. Thus, in a scene containing a green apple laid upon a towel and some other non-apple objects, if a speaker is trying to draw attention to the apple, stating *the apple* is sufficient, as mentioning other properties of the apple (such as its color or location) is not needed (and therefore redundant) in terms of the communicative goal.

However, speakers frequently violate the Maxim of Quantity, introducing redundant information into a sizable portion of their utterances (e.g., Deutsch & Pechmann, 1982; Engelhardt et al., 2006), including when interacting with artificial entities such as robots (Saryazdi et al., 2021). While redundancy is common, it remains unclear whether redundancy usage in language benefits or detracts from comprehension. There is evidence that redundancy may confuse comprehenders, who may interpret the extra information as relevant (Levinson, 2000). For example, *the green apple* could suggest the presence of multiple apples of different colors in the scene, leading to confusion when there is only one apple in the scene (e.g., Sedivy et al., 1999). Indeed, Engelhardt et al. (2011) demonstrated that comprehenders experienced confusion (as reflected in behavioral and electrophysiological measures) when hearing redundant utterances. However, other studies have found evidence that certain attributes, such as color, do not cause confusion or slow down comprehension (Sedivy, 2003; Fukumura & Carminati, 2022; Fukumura & van Gompel, 2017).

In fact, redundancy usage has also been found to aid in language comprehension. Deutsch and Pechmann (1982) argue that redundancy aids search efficiency (and thus comprehension) when the communicative goal is object identification. Indeed, speakers often use redundant color adjectives in object naming when the scene contains variously colored objects, indicating that redundancy helps comprehenders identify the relevant object (Rubio-Fernández, 2016). In addition, redundancy is more likely to be used when the redundantly mentioned attribute is sufficient on its own to correctly specify a referent (i.e., when the color of the referent is different

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

than the rest of the other objects; van Gompel et al., 2019) and in tasks where accuracy in comprehension is important (Arts et al., 2011), suggesting that speakers can use redundancies to facilitate communication. Likewise, using color redundantly has been found to facilitate visual referent search, with comprehenders finding a referent more quickly when its color was also given (e.g., *blue star* vs. *the star* when there was only one star in the scene; Rubio-Fernández, 2021). Similarly, listeners fixated on the referred objects to a greater extent (suggesting facilitated comprehension) in response to descriptions with redundant color information produced by a robot (especially when color can be used to rule out all other possibilities, with state modifiers being found to impair comprehension; Saryazdi et al., 2022). Finally, listeners deem utterances with redundancy as no less effective than those without (Engelhardt et al., 2006). Therefore, redundancy usage may facilitate communication by providing additional information to ensure that a listener has sufficient information to achieve the communicative goal, as information can be missed or misinterpreted in communication due to various reasons such as lack of attention, environmental distractors, and many other factors. In such scenarios, redundant information may be interpreted by the listener not as unnecessary information but as useful additional information and thus may not lead to confusion.

Thus, the usage or avoidance of redundancies might both be a goal-oriented choice by the speaker to aid comprehension for the comprehender. Indeed, speakers have been found to utilize redundancies in this manner, in that they modify the rates of their redundancy usage based on the extent to which a redundant word can help further identify a referent (Rubio-Fernández, 2021). In addition, because redundancy may or may not be typically expected, when a speaker opts for redundant or non-redundant utterances, it may indicate to their conversation partner that they prefer this style of communication, prompting the partner to employ redundancy as a strategy to facilitate understanding and ease of communication with their interlocutor.

Linguistic alignment

Linguistic redundancy may also emerge due to linguistic alignment, which refers to the tendency for interlocutors in a dialogue to imitate each other's linguistic behaviors at various levels (Pickering & Garrod, 2004). People have been found to mimic their interlocutor's speech rate and accent (e.g., Giles et al., 1991), phonetics (Pardo, 2006), lexical choices (e.g., Brennan, 1996), sentence structures (Branigan et al., 2000), conceptualizations of objects and scenarios (Garrod & Anderson, 1987), and even extra-linguistic aspects such as facial expressions (Bavelas et al., 1986; Dimberg et al., 2000), body posture (Tia et al., 2011), and speech gestures (Goldin-Meadow & Alibali, 2013).

According to the interactive alignment account (Pickering & Garrod, 2004), various linguistic representations used in dialogue tend to align automatically. This process results in a broad alignment ranging from low-level phonetics to high-level situational representations of the conversation. Under this theory, the encounter of a specific linguistic feature (e.g., phonetic realization, lexical choice) activates the corresponding representation. This residual activation increases the probability that

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

the same linguistic feature will be produced in subsequent interactions (Pickering & Branigan, 1998).

Linguistic alignment might also be driven by goal-oriented processes. Recent evidence suggests that the degree of alignment can vary depending on the perceived linguistic competence of the interlocutor. For example, speakers tend to more often align with the interlocutor and reuse their previous expressions when conversing with a non-native interlocutor than with a native interlocutor (Cai et al., 2021; Suffill et al., 2021). This pattern suggests that speakers consider certain characteristics (like linguistic competency and age) when crafting their utterances, a process referred to as *interlocutor modeling* (Cai et al., 2021; for a review, see Wu & Cai, 2024); in particular, when speakers perceive their interlocutor to be limited in linguistic competence, they are more likely to reuse their interlocutor's prior lexical choices to maximize communicative success. In further support of interlocutor modeling, speakers tend to perceive computers to have limited linguistic capacity and are also more likely to lexically align with a computer interlocutor than a human interlocutor (Branigan et al., 2004). Linguistic alignment as a form of goal-oriented process also concurs with the tenets of communication accommodation theory (Giles, 1973; see Giles et al., 2023 for a review), which states that interlocutors tend to converge in their linguistic patterns due in part to meeting the expectations of the interaction (Giles, 2008), and this convergence has been found to occur when interacting linguistically with artificial entities as well (e.g., Cirillo et al., 2022; Shen & Wang, 2023). Overall, interlocutor modeling should be a core part of communication accommodation theory, in that in order to tailor linguistic productions to a specific interlocutor, a speaker must first take notice of and create a mental model of the particular characteristics of their interlocutor, and then from this information, a speaker can then choose what linguistic aspects of their interlocutor to converge with.

Therefore, linguistic redundancies may arise as a result of alignment with an interlocutor in both linguistic and extra-linguistic contexts (Loy & Smith, 2021). If linguistic alignment (additionally) mirrors a goal-oriented process like interlocutor modeling, redundancy alignment might vary among different interlocutors. This variability in redundancy alignment may stem from people using more redundancies towards interlocutors with lesser linguistic capacity (such as computers), and especially if these interlocutors use redundancies in their own productions, as redundancies can provide additional linguistic cues to increase comprehensibility and ultimately communicative success (e.g., Rubio-Fernández, 2021). In this paper, we specifically examine whether speakers use (or do not use) redundancies as a result of alignment with an interlocutor and if the occurrence of such redundancy alignment differs when interacting with a computer versus a human interlocutor.

Linguistic redundancy alignment in human-computer interaction and human-human interaction

Naturally, computers and humans are very different entities, and indeed, differences have been found between human-computer interaction (HCI) and human-human interaction (HHI). For example, in typed conversations where people are told they are interacting with either a computer or a human interlocutor (but in real experiments, interlocutor responses are often scripted), people use fewer words and

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

less interpersonal language when addressing computers compared to humans (Shechtman & Horowitz, 2003). Similarly, speakers use fewer adjectives to describe a target object to a computer compared to a human; presumably, speakers perceive computers as being less competent in speech segmentation, thus being more likely to use shorter utterances to aid the computer in comprehension (Bannon et al., 2020). Furthermore, people tend to speak in a simpler and clearer way towards computers compared to humans by way of slower speech, hyper-articulation, fewer disfluencies, and more mimicry of words a computer uses (e.g., Oviatt, 1995; Bell, Gustafson, & Heldner, 2003; Stent, Huffman, & Brennan, 2008; Shen & Wang, 2023). Likewise, at the neural level, people are less surprised when LLMs make mistakes related to meaning, as shown by reduced brain activity upon encountering an anomalous word (but are more surprised when LLMs make grammatical mistakes; Rao et al., 2024).

However, HCI and HCI show striking similarities in certain areas. When engaging with virtual agents, people cooperate and communicate in ways akin to HHI (Parise et al., 1999; Krämer, 2005). These virtual agents are regarded by people as social entities (Krämer, 2005), can readily attract attention (Dehn & Van Mulken, 2000), induce socially desirable behavior (Sproull et al., 1996), and are treated with the same spatial usage rules and politeness norms (Kopp et al., 2005). People also use language in similar ways with virtual agents as humans (e.g., Saryazdi et al., 2021 van Lierop, Goudbeek, & Krahmer, 2012; Bergmann, Branigan, & Kopp, 2015) and interact with artificial intelligence assistants (i.e., Amazon's Alexa) in similar ways as compared with interacting with another human (Cohn & Zellou, 2021; Mengesha et al., 2021). In addition, people perceive artificial intelligence systems and robots as social actors and not merely mechanical tools (Bartneck et al., 2009; Groom et al., 2011), and people consistently view these systems as having high competency when acting independently from humans (McKee et al. 2023) and, in some circumstances, even moral standing (Malle et al., 2015, 2019). These findings suggest that, due to humans' inherent sociality, the social rules applied in HHI are also used unconsciously in HCI (Nass et al., 1997) and that people view these agents as having human-like perceptual and linguistic abilities (Saryazdi et al., 2021). Thus, due to the rapid development of computers and artificial intelligence capabilities, people may not perceive computers to be as different from humans as in the past, as computers now possess much more advanced linguistic capabilities in more humanistic realms such as pragmatic abilities (Barattieri di San Pietro et al., 2023).

Overall, technological advances entail that HCI will continue to more closely approximate HHI. However, even in an era whereby artificial intelligence possesses high linguistic competencies, people still seem to behave differently in subtle ways in HCI. One intriguing behavior that might share broad similarities but reveal key differences between HCI and HHI is linguistic alignment. This phenomenon is observed in HCI at various levels, including in phonetics (Gessinger et al., 2019), the lexicon (Brennan, 1996), and syntax (Branigan et al., 2003). The automatic mechanism of linguistic alignment should also function in HCI, as it is triggered by merely processing a linguistic element, regardless of the interlocutor's identity. Therefore, interacting with a computer should induce similar alignment as with a human. However, differences may arise from interlocutor modeling due to the significant contrast in linguistic competence between computer and human

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

interlocutors (see Shen & Wang 2023, for recent evidence for lexical alignment differences in HCI compared to HHI). People may perceive computers as linguistically less competent, leading to more alignment in HCI to facilitate the interaction (Branigan et al., 2004; Branigan et al., 2011; Branigan et al., 2003). Moreover, differences among types of computers can modulate alignment. Pearson et al. (2006) found that participants aligned more with the lexical choices of a basic computer than an advanced one, even though they interacted with identical pre-scripted responses.

In this paper, we explore whether speakers align in linguistic redundancy with their interlocutor and whether such alignment, if any, differs between computer and human interlocutors. This form of linguistic alignment is expected to occur with both computer and human interlocutors (i.e., is interlocutor-independent) according to the interactive alignment model, whereby alignment is predicted to happen across a broad range of linguistic levels, including higher-order levels such as the pragmatic alignment of redundancy usage. In addition, redundancy alignment may result from goal-oriented language use. That is, speakers may anticipate that their counterparts understand redundancy in the same way they themselves employ it. Consequently, speakers may replicate this redundancy to enhance communication effectiveness, particularly when dealing with less skilled conversational partners such as computers. Speakers may also view redundancy usage as a strategy that aids a comprehender in picking out a referent (Rubio-Fernández, 2021). Thus, speakers may use more redundancies (and more redundancy alignment) towards computers via interlocutor modeling as a goal-directed strategy to help this less competent interlocutor in comprehension, especially if the computer also uses redundancies.

Furthermore, an important factor that may influence linguistic alignment via interlocutor modeling is theory of mind (ToM) abilities, that being the capacity to understand others' mental states such as beliefs, knowledge, feelings, etc. (Premack & Woodruff, 1978). ToM is thought to be crucial for successful social interaction to occur (e.g., Tooby & Cosmides, 1995), as making inferences about others provides useful information on how best to interact with others. Thus, superior ToM abilities may enable speakers to make more accurate inferences about their interlocutors, such as inferences on the linguistic competency of their interlocutor, thereby allowing more effective alignment with the interlocutor's linguistic tendencies. Besides making inferences about humans, people have been found to use ToM skills when interacting with artificial entities, such as inferring the perspectives of robots (Wahn et al., 2023; Zhao & Malle, 2022). However, taking notice of and understanding the supposed mental state and abilities (or lack thereof) of artificial entities such as a computer interlocutor may require greater ToM abilities than when inferring these qualities about humans, as people in general have more knowledge about humans compared to artificial entities such as computers (Epley et al., 2007). Therefore, we predict that higher ToM abilities may increase the difference in alignment rates towards computers compared to humans due to increased realization and attention given towards specific properties of computer interlocutors (i.e., reduced competency and how this contrasts with humans' greater competency) that necessitate more goal-directed language production.

Finally, we examine the time course of redundancy usage and alignment (if any). If redundancy usage is effortful, then we should expect redundancies to decrease as

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

the interaction progresses due to speaker fatigue. The Maxim of Quantity (Grice, 1975) also predicts that redundancy usage should decrease over time due to the inclination of speakers to make their utterances more efficient. In contrast, the interactive alignment account (Pickering & Garrod, 2004) predicts that redundancy usage may increase over time towards interlocutors who use redundancies, as this account theorizes that alignment in general increases over time. Thus, if alignment increases with redundant interlocutors over time, then redundancy usage will increase as well. We test these predictions using separate analyses with trial number as a predictor of redundancy usage.

The current study

To investigate redundancy alignment between an interlocutor perceived as linguistically more competent (human) versus less competent (computer), we employed a joint picture matching and naming task across three experiments. Both the "human" and "computer" interlocutors were, in reality, pre-scripted utterances to ensure identical behavior. In the task, participants and their interlocutors alternated in matching and naming pictures. In a matching trial, participants were presented with four shapes in a scene and then received a description from their interlocutor. Their task was to select the corresponding shape based on the description. In a naming trial, they were tasked with describing one shape in the scene for their interlocutor to identify.

Importantly, the interlocutor either consistently used redundancies or avoided them completely. Three attributes of shapes were included across the interlocutor responses in the experiments (color, size, and shading, where shading refers to light and dark variants of a color, e.g., light and dark green) in order to provide a range of information to the participants that was either redundant or non-redundant. Some attributes, such as color, tend to be used redundantly to a greater extent than other attributes, such as size (e.g., Engelhardt & Ferreira, 2014; Pechmann, 1989); thus, this increased prevalence may cause attributes such as color to be seen as less redundant; hence, the need for a range of attributes across our experiments. Overall, this study examines the following research questions: 1) Do speakers align with the redundancy usage patterns of their interlocutor? 2) If so, does this alignment come about from an interlocutor-independent automatic mechanism and/or a goal-directed mechanism due to interlocutor modeling and perceptions of interlocutor competency? 3) Does ToM modulate rates of redundancy alignment via effects on interlocutor modeling? 4) What is the nature of the time course of redundancy alignment? With regard to these questions, we hypothesized that participants would mirror their interlocutor's redundancy usage, leading to an increase in redundant responses when interacting with a redundant interlocutor as compared to a non-redundant interlocutor. Furthermore, we expected this redundancy alignment to be more prominent with a linguistically less competent interlocutor (i.e., the computer interlocutor). To further delve into the relationship between linguistic alignment of redundancy usage and other variables, we also evaluated participants' ToM abilities, their perceptions of the interlocutor's competency, and the time course of this alignment (if any). For assessing ToM abilities, we used the "Reading the Mind in the Eyes" Test (Baron-Cohen et al., 2001) as a measure of ToM, as this measures the

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

ability to perceive the thoughts and feelings of others based on little information, which relates to making inferences in general about the traits of one's interlocutor. Overall, the results of these assessments will provide us with more insights into how redundancy alignment functions in different interaction contexts and how it might be influenced by cognitive abilities, interlocutor perceptions, and time.

A few previous studies (Loy & Smith, 2021; Goudbeek & Krahmer, 2012) have found evidence of redundancy alignment. However, participants in these studies had the opportunity to produce or not produce redundancies in various filler trials, but the utterances in these trials were not examined, making the extent of redundancy alignment unclear. The filler trials also may have influenced the redundancy usage behavior of the target trial by priming redundancy usage with descriptions that may be considered to include redundancy by the participants (e.g., the description *the woman who looks angry* and *the angry-looking woman* to describe a picture of an angry woman), further obfuscating the exact nature of how redundancy alignment operated in these studies. Thus, the current study expands on this research by removing these potentially confounding filler trials to examine whether redundancy alignment occurs directly after comprehending a redundant utterance.

Experiment 1

Methods

Participants

To determine an appropriate sample size for Experiment 1 (as well as the other experiments in this study), a power analysis was conducted with GPower 3.1.9.6 (Faul et al., 2007). This power analysis found that for a 2-by-2 repeated measures design with medium effect sizes, 80% power at an alpha of 0.05 is achieved at 64 participants. Thus, for Experiment 1 (and all other experiments), we set out to obtain sample sizes exceeding this.

A total of 114 native English speakers who were British nationals residing in the United Kingdom (mean age = 36.03; 46 male, 68 female; 101 white, 4 black, 7 Asian, 2 mixed) were recruited using the online participant recruitment platform Prolific (https://www.prolific.co/), with 21 participants removed from the analyses due to the screening criteria (see results section below for more details). All participants in this experiment (as well as all other experiments in this study) gave their informed consent to participate, and this study was given ethical approval by the Survey and Behavioural Research Ethics Committee at the Chinese University of Hong Kong (ethics code SBRE-22-0464).

Materials

The materials for the main experimental trials consisted of 64 scenes of four shapes arranged in a 2×2 grid (see Figure 1). Each scene consisted of a target shape (the shape to be matched or named) and three filler shapes. The position of the target shape was balanced across the scenes, such that the target shape was found equally in the top/bottom and left/right grid positions. The shapes varied and were balanced in size (large, small), color (blue, red), and shape (circle, triangle, square, ellipse,

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

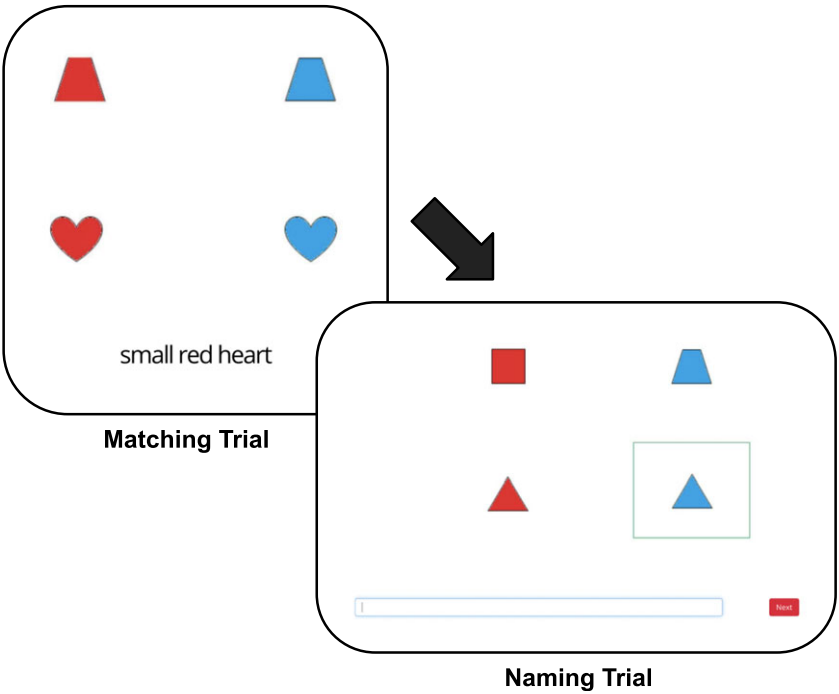

Figure 1. Example of a matching and naming trial. For the matching trial, the interlocutor gave a typed description of a shape, with participants needing to click on the described shape. For the naming trials, the target shape was outlined by a green border and participants described the shape via typing to their interlocutor. Note that the matching trial shows the redundant interlocutor condition, with *small* being redundant. The corresponding non-redundant interlocutor condition consists of the utterance *red heart*.

diamond, heart, trapezoid, pentagon). Two lists of 32 scenes were used, with each list consisting of 16 scenes where one of the filler shapes was the same as the target shape (e.g., the target being a big blue square and one filler being a small blue square) and 16 scenes without any fillers being the same shape as the target. Thus, if a target shape was presented together with a same-shape filler (i.e., an adjective was needed to single out the target) in one list, it was presented with fillers of different shapes (i.e., no adjective was needed to single out the target) in the other. The scenes from each list were used either in the matching or naming trials (counterbalanced between participants; i.e., in a Latin square design, that being if a scene with a small blue triangle as the target shape is used in a matching trial in one list, then this scene would be used in a naming trial in the other list).

The ToM task consisted of an online adaptation of the "Reading the Mind in the Eyes Test" (36 items; Baron-Cohen et al., 2001). This online version was designed to be as similar as possible to the original in-person version, whereby participants saw pictures of human eyes along with four emotion words (these word choices did not vary within each item, entailing that our items were the exact same as the original). One of the words correctly described the emotion of the person depicted, while the other three were incorrect choices. Participants were instructed to choose the word they thought best described the emotion of the person depicted. While online

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

adaptations may differ in reliability from the in-person equivalent, our results on this test when pooling data across all of the experiments (mean items correct = 25.5, SD = 4.30) are quite similar to the findings of the in-person test (mean items correct = 26.2, SD = 3.6; from Baron-Cohen et al.), suggesting that this online implementation is comparable to the in-person test. As Baron-Cohen et al. did not report any reliability metrics, we used the Cronbach's alpha to test for internal consistency in our data. This metric was found to be 0.616, suggesting that this measure has questionable reliability.

Procedure

The experiment was run online using the experiment-building platform Gorilla (https://www.app.gorilla.sc/). Participants gave their consent and then were told that they would be playing a picture matching and naming game with either a computer or human partner. Afterwards, a screen appeared informing participants that the study was waiting for another participant to conduct the experiment as their partner (the human condition, with this screen lasting for around 40 seconds to simulate another human partner being found and joining the study) or that the computer program was being loaded to conduct the experiment as their partner (the computer condition, with this screen lasting for around 7 seconds to simulate loading a computer program). The main trials then commenced, in which participants took turns matching and naming shapes with their supposed partner. In matching trials, participants received a typed description from their partner and then needed to click on the corresponding picture. In naming trials, participants typed a description into a text box under the scene. After the main trials, participants completed the "Reading the Mind in the Eyes Test" and then filled out a post-experiment questionnaire. The post-experiment questionnaire asked participants to give information regarding 1) if they thought their partner was a computer or a human, 2) how competent they thought their partner was in the experimental task on a 7-point Likert scale, and 3) if they had any other comments about the experiment. Participants were paid 1.5 British pounds after finishing the experiment, which took around 15 minutes to complete.

Data coding

The typed responses for the naming trials were coded as either redundant or non-redundant. A redundant coding was given to a response if it included at least one redundancy (e.g., the response *blue square* would be redundant if only one square was present in the scene, as in this case, the descriptor *blue* would not be needed). A non-redundant coding was given to a response if it included no redundancies (e.g., the response *square* for the previous example). Responses that were under-informative were excluded from further analyses (e.g., the response *square* in a scene with two squares).

Statistical analyses

In Experiment 1 (as well as in Experiments 2 and 3), binomial logit mixed effects (LME) models were used as the primary analyses to examine the role of interlocutor

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

(human vs. computer) and usage (redundant vs. non-redundant) on whether participants used redundancies or not in their responses. In order to establish the random effects structure for these models, a data-driven approach via forward model comparisons was used (e.g., Matuschek et al., 2017; Bates et al., 2015). This process starts by creating an intercept-only LME model (i.e., containing only subject and item random intercepts but no random slopes). Then, this random intercept-only model is compared to a model with the inclusion of an additional random slope (using the *anova* function in R). A significant *p*-value from the model comparison (with alpha set to 0.2 instead of the typical 0.05 to achieve a balance between an overly basic and overly maximal random effects structure; Matuschek et al., 2017) suggests that the inclusion of the additional random slope improves the model fit over the intercept-only model and entails the inclusion of this random slope. This model comparison process is repeated, with a new random slope added to the current best model, and then this new augmented model is compared to the current best model until no additional random slopes improve model fit. Thus, this model selection process is designed to find an optimal balance of random effects to include in a statistical model. The steps of each model comparison process, as well as the data and analytical scripts for all experiments in this study, can be found at https://osf.io/2kzn9/.

Results

Among the 114 participants, 19 participants did not believe that their interlocutor was a computer/human when told that they would be conducting the experiment with a computer/human, respectively, and their responses were removed from further analyses. Two further participants incorrectly named the target shape more than one-third of the time and were excluded from further analyses. Of the remaining 3041 responses, 182 were incorrect (under-informative or incorrect descriptions of the target shape) and were removed from further analyses, leaving a total of 2859 usable trials (52 and 41 participants in the computer/human conditions with 1598 and 1261 trials, respectively). Participants on average produced redundant and non-redundant descriptions in 35.4% and 64.6% of their responses, respectively.

We first applied a binomial LME model on trial-level responses (i.e., redundant and non-redundant, with the latter being the baseline level), using usage (redundant = −1 vs. non-redundant = 1 interlocutor; same coding in all following models with usage) and interlocutor (human = −1 vs. computer = 1; same coding in all following models with interlocutor) as interacting predictors.

As detailed in Table 1, the significant intercept indicated that there were fewer redundant than non-redundant responses (35.4% vs. 64.6%). The significant effect of usage suggested that participants were more likely to produce redundant descriptions when addressing a redundant interlocutor (55.7%) than a non-redundant interlocutor (16.3%; see Figure 2). The main effect of interlocutor was also significant, suggesting that participants increased their redundancy usage when interacting with a computer (39.6%) compared to a human (30.1%). The interaction between usage and interlocutor was not significant, indicating that the effect of usage does not differ between the human and computer interlocutors.

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

Table 1. LME results with usage and interlocutor (conditional R^2 = .430, marginal R^2 = .217)

Fixed Effects	*β*	*SE*	*z*	*p*
Intercept	−0.91	0.14	−6.31	<.001
Usage	−1.09	0.10	−10.78	<.001
Interlocutor	0.24	0.11	2.13	.033
Usage: Interlocutor	0.11	0.10	1.13	.259
Random Effects	Estimate (explained variance/standard deviation/correlation)			
Subjects: (intercept)	0.75/0.87			
Subjects: Usage	0.21/0.46/0.27			
Items: (intercept)	0.26/0.51			

Note: This model including participants who did not believe the interlocutor manipulation produced the same results except for the main effect of interlocutor, with a significant intercept and main effect of usage, and non-significant interlocutor and interaction effects.

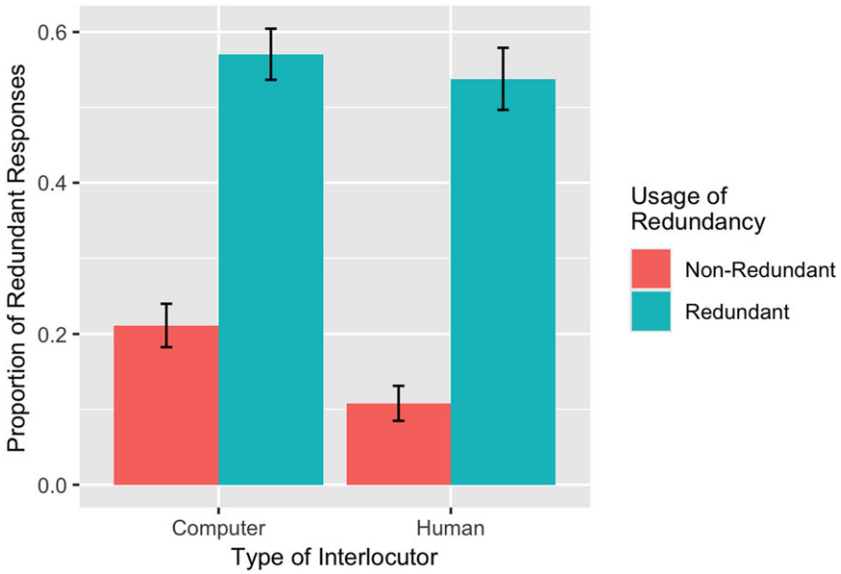

Figure 2. Proportion of redundant responses for usage and interlocutor (error bars show the 95% CIs).

In order to explore how redundancy usage progressed across the trials, a further LME model (Table 2) was built with usage and trial (logged trial number; this variable was logged in order to aid in model convergence) as interacting predictors. Usage was found again to have a significant main effect, and more crucially, there was a significant main effect of trial, suggesting that participants produced fewer redundant descriptions as time progressed. The interaction between usage and trial was also significant (see the appendix for all of the figures related to the trial analyses), with separate analyses indicating that the effect of trial was stronger in the

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

Table 2. LME results with usage and trial (conditional $R^2 = .455$, marginal $R^2 = .220$)

Fixed Effects	*β*	*SE*	*z*	*p*
Intercept	0.48	0.21	2.31	.021
Usage	−0.48	0.16	−3.00	.003
Trial	−0.55	0.06	−9.40	<.001
Usage: Trial	−0.21	0.06	−3.69	<.001
Random Effects	Estimate (explained variance/standard deviation/correlation)			
Subjects: (intercept)	1.15/1.07			
Items: (intercept)	0.27/0.52			

Note: This model including participants who did not believe the interlocutor manipulation produced the same results. This model with trial as a non-logged predictor (and all the models in this study including trial) produced the same results.

Table 3. LME results with interlocutor, ToM, and competency (conditional $R^2 = .337$, marginal $R^2 = .020$)

Fixed Effects	*β*	*SE*	*z*	*p*
Intercept	2.28	1.24	1.83	.067
Interlocutor	−1.16	0.74	−1.57	.116
ToM	0.23	1.08	0.21	.830
Competency	−0.20	.178	−1.13	.258
Interlocutor: ToM	1.33	1.05	1.27	.203
Random Effects	Estimate (explained variance/standard deviation/correlation)			
Subjects: (intercept)	1.41/1.19			
Items: (intercept)	0.16/0.40			

Note: This model including participants who did not believe the interlocutor manipulation produced the same results.

non-redundant condition ($\beta = -0.74$, $SE = 0.09$, $z = -8.36$, $p < .001$) compared to the redundant condition ($\beta = -0.33$, $SE = 0.08$, $z = -4.22$, $p < .001$).

Another LME model (Table 3) was built to explore if ToM abilities and the perceived competence of the interlocutor affect redundancy alignment (i.e., aligned and non-aligned, with the latter being the baseline level; this coding was used for all subsequent analyses that use redundancy alignment as the dependent variable), with interlocutor and ToM as interacting predictors and competency as a non-interacting predictor (in order to aid in model convergence). None of the effects involving these predictors were significant. These results suggest that increased ToM abilities do not aid in establishing a more definite interlocutor model of a human compared to a computer, or if so, that these differentiations do not affect redundancy alignment in language production. Likewise, perceived interlocutor competency (as well as whether the interlocutor is a human or computer) does not seem to affect redundancy alignment.

A Wilcoxon rank-sum test using the means of the perceived competency data for human and computer interlocutors was done to explore if participants viewed

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

humans or computers as more competent in the experimental task. This test was used over other similar tests, as perceived competency is not normally distributed for both the human and computer conditions ($W = 0.47$ and 0.59, respectively, and ps $<.001$), while homogeneity of variances between these conditions was met ($F(1) = 0.844$, $p = .361$). No significant difference was found for this test ($W = 949.5$, $p = 0.227$), suggesting that humans (6.71) and computers (6.56) were not perceived as having different competency levels.

Discussion

Experiment 1 had participants play a picture matching and naming game with an interlocutor who either always or never included redundancies in their responses, with the interlocutor being either a computer or a human. People used more redundancies when interacting with a redundant interlocutor compared to a non-redundant interlocutor, indicating that speakers align to the redundancy usage patterns of their interlocutor. More redundancies in general were also used towards computers than humans. Redundancy usage also decreased over time, suggesting that redundancy usage is effortful and that utterances become more efficient with increased exposure to the linguistic task at hand. Redundancy usage decreased at a faster rate with non-redundant interlocutors compared with redundant interlocutors, entailing that people align to their interlocutor's redundancy usage behavior at the initial stages of interaction and that this alignment strengthens over time with non-redundant interlocutors but weakens with redundant interlocutors. The type of interlocutor, ToM, and the perceived competency of the interlocutor do not seem to affect redundancy alignment. We did not observe a difference in perceived linguistic competence between computer and human interlocutors, though as expected computer interlocutors were numerically rated as linguistically less competent than human interlocutors.

Experiment 2

While the results of Experiment 1 suggest that people align to the redundancy usage of their interlocutor, this alignment could also be merely due to lexical priming. That is, as the same shape properties were used between the matching and naming trials (i.e., red/blue and big/small), people interacting with redundant interlocutors may have been primed to produce the redundant adjectives they heard from the matching trials in the naming trials. For instance, participants might produce *big triangle* after hearing *big square*, not because they were primed to be redundant but because they tended to reuse the adjective *big* as a result of lexical priming. Experiment 2 addresses this issue by implementing the same tasks as in Experiment 1 but using different shape properties between the matching and naming trials so that there is no lexical overlap in the shape properties between the matching and naming trials, and hence no opportunity for lexical priming to occur.

Methods

A total of 114 participants from the same population as Experiment 1 (mean age = 34.73; 31 male, 83 female; 103 white, 3 black, 6 Asian, 2 mixed) were

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

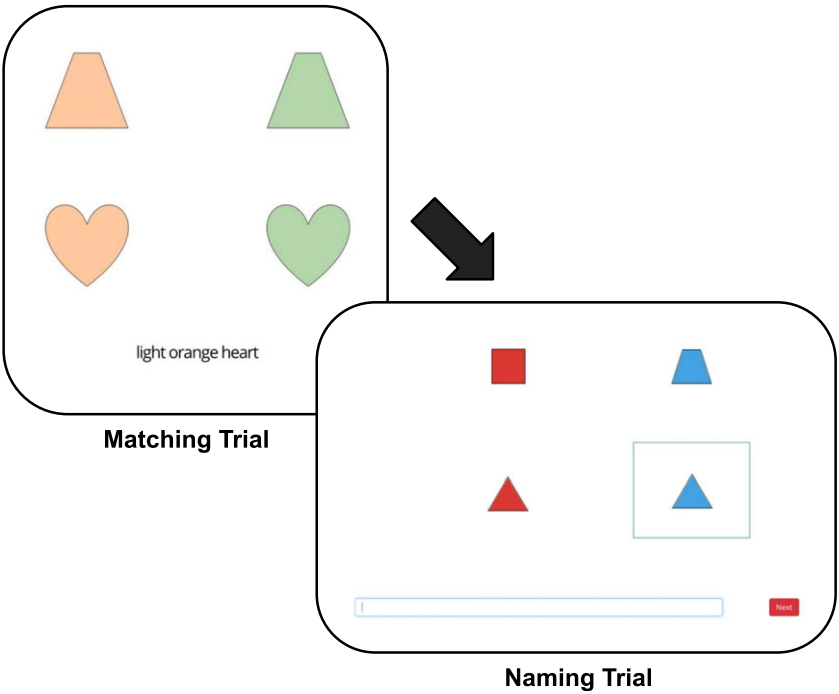

Figure 3. Example of a matching and naming scene for Experiment 2. The matching trials consist of different color and shading features compared to the naming trials.

recruited (none of these participants took part in Experiment 1), with 23 participants removed from the analyses due to the screening criteria. The materials for the main experimental trials were the same as in Experiment 1, except that the shape features were changed as follows for the matching trials: the colors blue/red from Experiment 1 were replaced with the colors orange/green, respectively, and the features big/small were replaced with the features light/dark colors, respectively (with all shapes being the same size; see Figure 3). These changes meant that no adjectives that appeared in the matching trials would be reused in the naming (e.g., the color orange may depict some shape attributes in the matching trials but would never be relevant for any shape attribute in the naming trials). The procedure and statistical analyses (as well as the specific rationale for the statistical analyses as detailed in section 2.1) were the same as those in Experiment 1.

Results

Of the 114 participants, 20 participants did not believe that their interlocutor was a computer/human when told that they would be conducting the experiment with a computer/human, respectively, and their responses were removed from further analyses. Three further participants incorrectly named the target shape more than one-third of the time and were excluded from further analyses, leaving a total of 2744 usable trials after removing 236 incorrect trials (55 and 36 participants in the computer/human conditions with 1658 and 1086 trials, respectively). Participants

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

Table 4. LME results with usage and interlocutor (conditional $R^2 = .548$, marginal $R^2 = .245$)

Fixed Effects	β	*SE*	*z*	*p*
Intercept	−0.85	0.18	−4.59	<.001
Usage	−1.35	0.17	−7.88	<.001
Interlocutor	0.31	0.16	1.92	.055
Usage: Interlocutor	0.27	0.16	1.67	.096
Random Effects	Estimate (explained variance/standard deviation/correlation)			
Subjects: (intercept)	1.85/1.36			
Items: (intercept)	0.25/0.50			
Items: Usage	0.10/0.32/−0.74			

Note: This model including participants who did not believe the interlocutor manipulation produced broadly the same results, with a significant intercept and main effect of usage and interlocutor, and a non-significant usage/interlocutor interaction.

on average produced redundant and non-redundant descriptions in 38.8% and 61.2% of their responses, respectively.

As in Experiment 1, a first analysis was done by building a binomial LME model with the redundant and non-redundant responses (using non-redundant responses as the baseline level) and including usage (redundant vs. non-redundant interlocutor) and interlocutor (human vs. computer interlocutor) as interacting predictors. As detailed in Table 4, there were fewer redundant (38.8%) than non-redundant (61.2%) responses, as indicated by the significant intercept (see Figure 4). Participants were more likely to produce redundant descriptions when addressing a redundant interlocutor (59.0%) than a non-redundant interlocutor (18.4%), as indicated by the significant main effect of usage. There was a marginally significant main effect of interlocutor, tentatively suggesting that participants increased their redundancy usage when interacting with computers (39.5%) compared with humans (37.8%). The interaction between usage and interlocutor was not significant, indicating that the effect of usage does not differ between the human and computer interlocutors.

In order to explore how redundancy usage progressed across the trials, a further LME model (Table 5) was built with usage and trial (logged trial number) as interacting predictors. Usage was found again to have a significant main effect, and more crucially, there was a significant main effect of trial, suggesting that participants produced fewer redundant descriptions as the experiment progressed. The interaction between usage and trial was also significant (see appendix for the figure), with separate analyses indicating that the effect of trial was stronger in the non-redundant condition ($\beta = -1.21$, $SE = 0.22$, $z = -5.61$, $p < .001$) compared to the redundant condition ($\beta = -0.40$, $SE = 0.17$, $z = -2.37$, $p = .018$).

Another LME model (Table 6; see appendix for figure) was built to explore if ToM abilities and the perceived competence of the interlocutor affect redundancy alignment, with interlocutor and ToM as interacting predictors and competency as a non-interacting predictor. There was a significant main effect of ToM, suggesting

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

Table 5. LME results with usage and trial (conditional R^2 = .600, marginal R^2 = .280)

Fixed Effects	*β*	*SE*	*z*	*p*
Intercept	0.88	0.25	3.57	<.001
Usage	−0.65	0.23	−2.82	.005
Trial	−0.68	0.06	−10.72	<.001
Usage: Trial	−0.29	0.06	−4.59	<.001
Random Effects	Estimate (explained variance/standard deviation/correlation)			
Subjects: (intercept)	2.25/1.50			
Items: (intercept)	0.30/0.55			
Items: Usage	0.08/0.27/ −0.84			

Note: This model including participants who did not believe the interlocutor manipulation produced the same results.

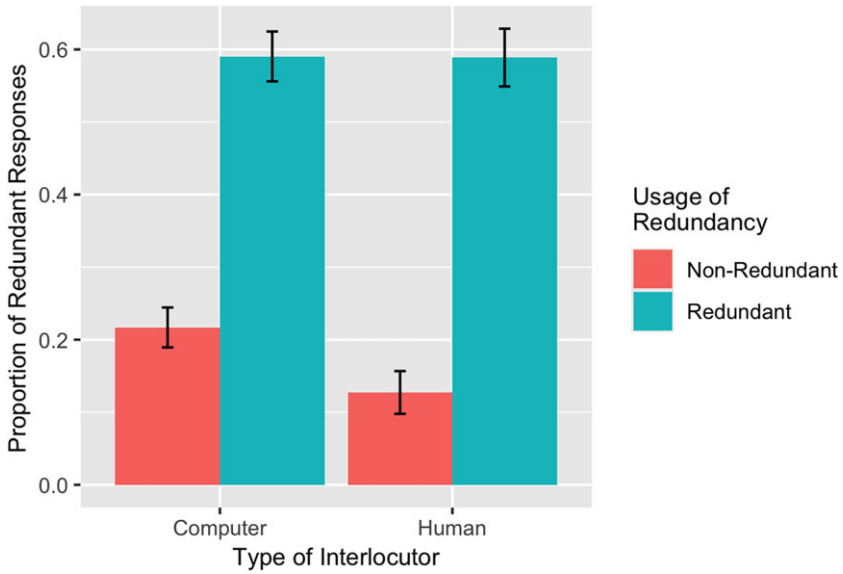

Figure 4. Proportion of redundant responses for usage and interlocutor (error bars show the 95% CIs).

that higher ToM scores were associated with less redundancy usage. However, the interaction between interlocutor and ToM was not significant, indicating that ToM does not aid in establishing an interlocutor model that affects alignment rates between human and computer interlocutors. There was a marginally significant effect of competency, indicating that interlocutors who were perceived as having higher competency were aligned to a greater extent. The main effect of interlocutor as well as the interaction between interlocutor and ToM were not significant, indicating that alignment rates do not differ between the human and computer interlocutors when taking into account ToM and competency, with ToM having similar effects across the human and computer interlocutors as well.

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

Table 6. LME results with interlocutor, ToM, and competency (conditional $R^2 = .435$, marginal $R^2 = .056$)

Fixed Effects	β	*SE*	*z*	*p*
Intercept	−3.32	1.40	−2.36	.018
Interlocutor	0.16	0.95	0.17	.863
ToM	3.47	1.31	2.66	.008
Competency	0.33	0.17	1.93	.054
Interlocutor: ToM	−0.47	1.30	−0.36	.717
Random Effects	Estimate (explained variance/standard deviation/correlation)			
Subjects: (intercept)	2.05/1.43			
Items: (intercept)	0.16/0.40			

Note: This model including participants who did not believe the interlocutor manipulation did not find any significant effects.

Table 7. LME results with usage and trial (conditional $R^2 = .413$, marginal $R^2 = .358$)

Fixed Effects	β	*SE*	*z*	*p*
Intercept	0.52	0.17	3.04	<.001
Usage	−0.77	0.15	−5.25	<.001
Trial	−0.41	0.06	−7.30	<.001
Usage: Trial	−0.23	0.06	−4.17	<.001
Random Effects	Estimate (explained variance/standard deviation/correlation)			
Subjects: (intercept)	0.07/0.27			
Items: (intercept)	0.23/0.48			

Note: This model including participants who did not believe the interlocutor manipulation produced the same results.

A Wilcoxon rank-sum test using the means of the perceived competency data for human and computer interlocutors was done to explore if participants viewed humans or computers as more competent in the experimental task. This test was used over other similar tests, as perceived competency is not normally distributed for both the human and computer conditions ($W = 0.62/\ 0.76$, respectively, and *p*s <.001), while homogeneity of variances between these conditions was met ($F(1) = 0.836$, $p = .357$). No significant difference was found for this test ($W = 837.0$, $p = 0.164$), suggesting that humans (6.44) and computers (6.25) were not perceived as having different competency levels.

Discussion

Experiment 2 differed from Experiment 1 in having no overlap in the target shape properties (e.g., size or color) between the matching and naming trials. As in Experiment 1, we found that participants used more redundancies when interacting with a redundant interlocutor compared to a non-redundant interlocutor, indicating

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

that speakers align to the redundancy usage patterns of their interlocutor. Crucially, this alignment could not have been due to lexical priming, as there was no lexical overlap between words that people heard to describe the shapes and words used in their productions. Redundancy usage also decreased over time (as in Experiment 1), suggesting again that redundancy usage is effortful and/or that utterances become more efficient with increased exposure to the linguistic task at hand. Redundancy usage decreased at a faster rate with non-redundant interlocutors compared with redundant interlocutors, entailing that people align to their interlocutor's (non-) redundancy usage at the initial stages of interaction and that this alignment strengthens over time with non-redundant interlocutors but weakens with redundant interlocutors. In contrast to Experiment 1, the type of interlocutor and ToM affected redundancy usage and alignment. Participants on average produced more redundant descriptions towards computers than humans, and higher ToM individuals also were found to align to a greater degree with the redundancy usage behavior of their interlocutor, which could be due to higher ToM allowing for one to take more notice of the specific linguistic behavior of their interlocutor and in turn copy this behavior in subsequent interaction. However, this effect of higher ToM and increased alignment was the same with human and computer interlocutors, suggesting that ToM does not affect the differentiating of interlocutor characteristics enough to impact any form of goal-directed redundancy alignment.

Experiment 3

In Experiments 1 and 2, the participants were merely told that they would be interacting with a computer or human interlocutor, and these interlocutors interacted with the participants via typed responses. Experiment 3 went a step further in distinguishing these two different types of interlocutors by having the computer and human interlocutors verbally produce their responses in a computerized and human voice, respectively. This was done to create a more apparent difference between these two interlocutors in order to further test if redundancy alignment is modulated by the type of interlocutor.

Methods

A total of 117 participants from the same population as Experiments 1 and 2 (mean age = 35.52; 39 male, 78 female; 105 white, 4 black, 4 Asian, 4 mixed) were recruited (none of these participants took part in Experiment 1 or 2), with 26 participants removed from the analyses due to the screening criteria. The materials for the main experimental trials were the same as in Experiment 1, except that the interlocutors' responses in the matching trials consisted of speech rather than typed script (with the interlocutor content being the same, and the participants in the matching trial still producing typed responses). The human speech stimuli were generated via the Google Text-to-Speech platform (https://cloud.google.com/text-to-speech), which produces naturalistic and consistent speech from text. This method of speech generation was chosen as it creates speech that is more consistent in terms of speech rate, volume, and vocal quality than layperson recordings. The computer speech stimuli were generated by applying a vocoder to the human speech

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

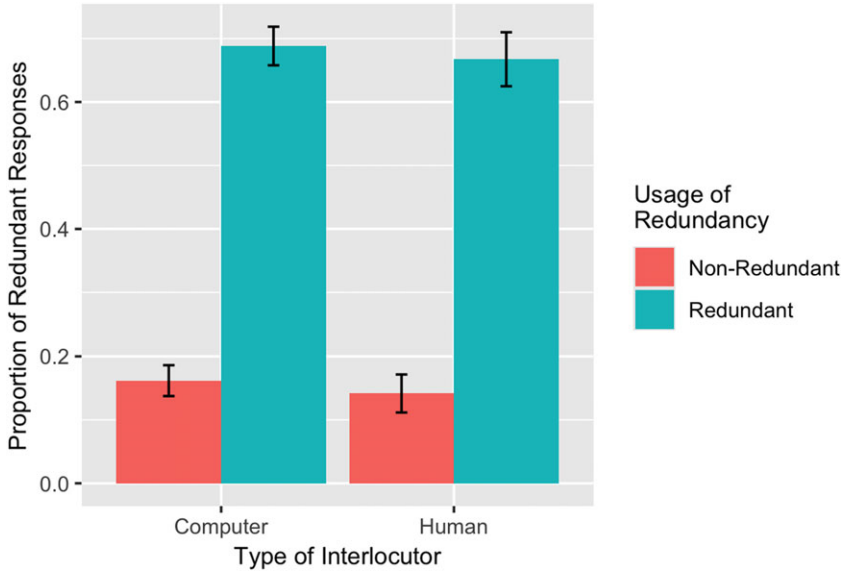

Figure 5. Proportion of redundant responses for usage and interlocutor (error bars show the 95% CIs).

stimuli, which transforms the speech to make it sound artificial and robotic. The procedure and statistical analyses (as well as the specific rational for the statistical analyses as detailed in section 2.1) were the same as in Experiment 1.

Results

Of the 117 participants, 26 participants did not believe that their interlocutor was a computer/human when told that they would be conducting the experiment with a computer/human, respectively, and their responses were removed from further analyses. Of the remaining 2981 responses, 204 were incorrect and were removed from further analyses, leaving a total of 2777 usable trials (58 and 33 participants in the computer and human interlocutor conditions, with 1779 and 998 trials, respectively). Participants on average produced redundant and non-redundant descriptions in 41.4% and 58.6% of their responses, respectively.

As in Experiments 1 and 2, a first analysis was done by building a binomial LME model with the redundant and non-redundant responses (using non-redundant responses as the baseline level) and including usage (redundant vs. non-redundant interlocutor) and interlocutor (human vs. computer interlocutor) as interacting predictors. As detailed in Table 6, there were fewer redundant (41.4%) than non-redundant (58.6%) responses, as indicated by the significant intercept (see Figure 5). Participants were more likely to produce redundant descriptions when addressing a redundant interlocutor (68.1%) than a non-redundant interlocutor (15.4%), as indicated by the significant main effect of usage. The main effect of interlocutor as well as the interaction between usage and interlocutor was not significant, indicating that redundancy rates do not differ between the human and computer interlocutors and that the effect of usage does not differ between these interlocutors.

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

Table 8. LME results with interlocutor, ToM, and competency (conditional $R^2 = .039$, marginal $R^2 = .003$)

Fixed Effects	*β*	*SE*	*z*	*p*
Intercept	0.58	0.40	1.45	.147
Interlocutor	−0.08	0.34	−0.23	.821
ToM	0.00	0.48	−0.01	.993
Competency	0.10	0.04	2.37	.018
Interlocutor: ToM	0.13	0.47	0.27	.791
Random Effects	Estimate (explained variance/standard deviation/correlation)			
Items: (intercept)	0.12/0.35			

Note: This model including participants who did not believe the interlocutor manipulation produced the same results.

Table 9. Summary of results across all experiments

Analysis	Experiment 1	Experiment 2	Experiment 3
Usage	<.001	<.001	<.001
Interlocutor	.033	.055	.267
Usage: Interlocutor	.259	.096	.627
ToM	.830	.008	.993
Interlocutor: ToM	.203	.717	.791
Competency	.258	.054	.018
Interlocutor competency	0.35	.360	.020
Trial	<.001	<.001	<.001
Usage: Trial	<.001	<.001	<.001

In order to explore how redundancy usage progressed across the trials, a further LME model (Table 7) was built with usage and trial (logged trial number) as interacting predictors. Usage was found again to have a significant main effect, and more crucially, there was a significant main effect of trial, suggesting that participants produced fewer redundant descriptions as time progressed. The interaction between usage and trial was also significant (see appendix for figure), with separate analyses indicating that the effect of trial was stronger in the non-redundant condition ($\beta = -0.62$, $SE = 0.08$, $z = -7.68$, $p < .001$) compared to the redundant condition ($\beta = -0.18$, $SE = 0.08$, $z = -2.34$, $p = .019$).

Another LME model (Table 8) was built to explore if ToM abilities and the perceived competence of the interlocutor affect redundancy alignment, with interlocutor and ToM as interacting predictors and competency as a non-interacting predictor. The main effect of competency was found to be significant, suggesting that participants aligned to a greater extent with interlocutors who were perceived higher in competency, with separate analyses indicating this result is driven by less redundancy usage towards non-redundant interlocutors who were perceived as more competent ($\beta = 0.21$, $SE = 0.06$, $z = 3.54$, $p < .001$; see Figure 6), with no

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

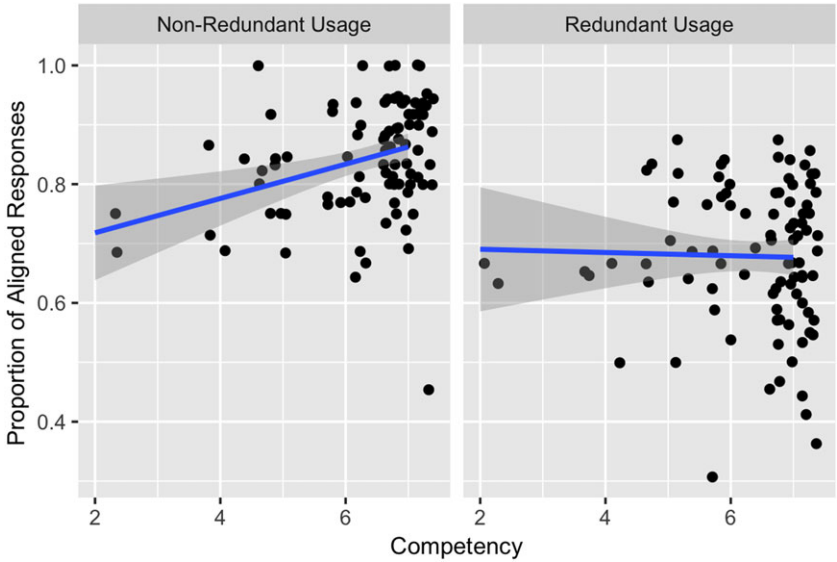

Figure 6. Proportion of aligned responses for competency, redundancy, and usage (shading shows the 95% CIs).

effect of competency towards redundant interlocutors ($\beta = 0.02$, $SE = 0.06$, $z = 0.31$, $p = .759$). The main effects of interlocutor and ToM, as well as the interaction between these factors, were not significant, indicating that redundancy alignment rates do not differ between the human and computer interlocutors and do not depend on ToM abilities.

A Kruskal-Wallis test using the means of the perceived competency data for human and computer interlocutors was done to explore if participants viewed humans or computers as more competent in the experimental task. This test was used over other similar tests, as perceived competency is not normally distributed for both the human and computer conditions ($W = 0.47/\ 0.59$, respectively, and ps $<.001$), and homogeneity of variances between these conditions was not met ($F(1) = 4.281$, $p = .041$). A significant difference was found for this test ($\chi^2(1) = 5.197$, $p = 0.023$), suggesting that humans (6.64) were perceived to be more competent compared to computers (6.16).

Discussion

Experiment 3 was the same as Experiment 1, except that participants received spoken (instead of typed) descriptions produced by their interlocutor in the matching trials. Similar to Experiments 1 and 2, people used more redundancies when interacting with a redundant interlocutor compared to a non-redundant interlocutor, indicating that speakers align to the redundancy usage patterns of their interlocutor. Redundancy usage also decreased over time (as in Experiments 1 and 2), suggesting again that redundancy usage is effortful and that utterances become more efficient with increased exposure to the linguistic task at hand. Redundancy usage

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

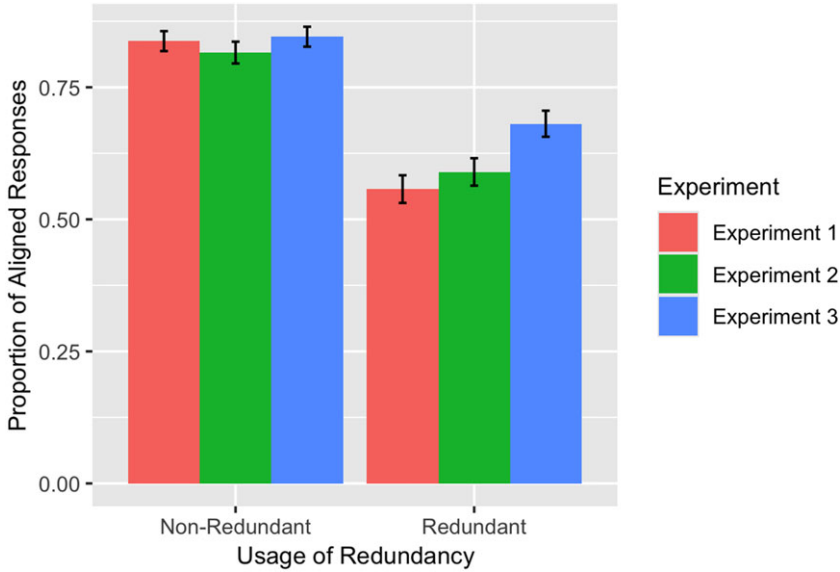

Figure 7. Proportion of aligned responses for usage and experiment (i.e., redundancy usage comparisons between Experiments 1, 2, and 3) (error bars shows the 95% CIs).

decreased at a faster rate with non-redundant interlocutors compared with redundant interlocutors, entailing that people align to their interlocutor's redundancy usage behavior at the initial stages of interaction and that this alignment strengthens over time with non-redundant interlocutors but weakens with redundant interlocutors. Competency also had an effect on redundancy alignment, with participants aligning more when they perceived a non-redundant interlocutor to be linguistically more competent. We also observed a significant difference in perceived linguistic competency, with lower competency ratings given to computer interlocutors than to human interlocutors.

Comparison across experiments

In order to compare redundancy alignment between the experiments, an LME model was built with experiment (experiments were dummy-coded with Experiment 1 as the baseline level) and usage as interacting predictors, with the dependent variable being redundancy alignment. The main effect of usage ($\beta = 0.76$, $SE = 0.07$, $z = 11.14$, $p < .001$) was found to be significant, indicating greater rates of redundancy alignment when interacting with a non-redundant interlocutor (83.3%) than a redundant interlocutor (60.9%). The significant interaction between usage and Experiment 3 suggests that alignment differences with redundant and non-redundant interlocutors differ in magnitude between Experiment 3 and Experiment 1 ($\beta = -0.26$, $SE = 0.08$, $z = -3.14$, $p = .002$; see Figure 7), with separate analyses revealing that there was more alignment in Experiment 3 compared to Experiment 1 with redundant interlocutors ($\beta = 0.65$, $SE = 0.14$,

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

$z = 4.69$, $p < .001$) but the same with non-redundant interlocutors ($\beta = 0.01$, $SE = 0.16$, $z = 0.05$, $p = 0.96$). Table 9 gives a summary of the results across the three experiments.

In addition, chi-squared tests were used in order to explore if under-informative response rates differed between computer and human interlocutors for each experiment. These tests all were non-significant ($p = .544$, 913, 897 for Experiments 1, 2, 3, respectively), indicating that people produce under-informative responses at similar rates towards computers and humans.

General discussion

The three experiments, along with the combined experimental analyses, demonstrate that speakers do produce redundancies and align with the higher-order linguistic behavior of redundancy usage (all experiments). This alignment is influenced by the perceived competency of the interlocutor (Experiments 2 and 3), ToM (Experiment 2), and the time course of the interaction (all experiments). These redundancy usage patterns were quite similar, regardless of whether the matching trials and naming trials shared the same shape descriptors (Experiment 1 vs. 2), while redundancy alignment was greater when people heard spoken descriptors compared to reading typed descriptors (Experiment 1 vs. 3). People in general used more redundancies with computers than humans as well (Experiments 1 and 2). In the following discussion, we further examine the role these factors play in redundancy usage.

Linguistic redundancy

Theoretical accounts of communication typically assert that speakers should eschew redundancies in their speech (Grice, 1975). Nevertheless, this investigation, in addition to preceding studies (Deutsch & Pechmann, 1982; Engelhardt et al., 2006), indicates that speakers frequently incorporate redundancies in their utterances. Although the bulk of responses were characterized by non-redundant descriptions, across the three experiments, redundancies featured in approximately one-third of responses, constituting a sizable minority (see Engelhardt et al., 2006, for a similar redundancy usage rate). Furthermore, redundancy usage was present even when speakers never heard redundant utterances from their interlocutor, suggesting that redundancy usage does not require prior encounters with redundant utterances. Therefore, while speakers typically adhere to the Maxim of Quantity, this principle can often be violated (in HCI as well, concurring with Saryazdi et al., 2021).

Why would speakers choose to produce redundancies, though? Redundancies, by their very nature, comprise superfluous information that may not be needed for achieving a communicative objective, such as referencing a particular object. Hence, producing redundancies should demand more effort than refraining from doing so, given that generating redundancies lengthens a speech utterance, thereby necessitating greater cognitive and physical exertion on both the speaker and the audience. It is plausible that speakers opt to include redundancies under the belief that they assist the comprehender in seeking out and identifying the referenced object by ensuring that the comprehender has additional information in case some

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

information is missed in comprehension (Deutsch & Pechmann, 1982; Rubio-Fernández, 2016, Rubio-Fernández, 2021; Saryazdi et al., 2022). In some scenes across the three experiments, redundantly mentioning the color or size of the target shape could reduce the number of possibly referred-to shapes by half, which could potentially aid the interlocutor in successfully identifying the target shape (e.g., Rubio-Fernández, 2016, 2021; Deutsch & Pechmann, 1982). Consequently, the use of redundancies appears to be a calculated strategy (alongside being driven by linguistic alignment) to enhance communication success, albeit at the expense of increased effort on the speaker's part.

Redundancy alignment in HCI and HHI

The increase in redundancy usage when interacting with a redundant interlocutor (both in HCI and HHI) is consistent with the interactive alignment model (Pickering & Garrod, 2004), which suggests that speakers align with their interlocutor's redundancy usage, leading to convergence towards the interlocutor's linguistic behavior. Primarily, this redundancy alignment appears to stem from an automatic and interlocutor-independent mechanism. This mechanism asserts that processing redundant utterances triggers residual activation of a redundant linguistic element, thereby amplifying the probability of future redundancy usage. The extent of this alignment was considerable; redundancies were employed in about 61% of responses when conversing with a redundant interlocutor, in contrast to only about 17% of responses when interacting with a non-redundant interlocutor. These findings indicate that people align with higher-order linguistic patterns in addition to lower-order patterns, such as in phonetic (Pardo, 2006), lexical (Brennan, 1996), and syntactic (Branigan et al., 2000) alignment. This is also consistent with the finding that people align with their interlocutor's usage of basic and superordinate terms (Cirillo et al., 2022), which can be considered another form of higher-order linguistic behavior. Note that this redundancy alignment seems to be independent of lexical alignment, as suggested by the comparison between Experiments 1 and 2 (and in agreement with Loy & Smith, 2021). This further supports the notion of the existence of a higher-order redundancy element, which differs from the mere activation of a lexical item. However, this contrasts with previous findings of lexical boost effects in syntactic alignment (e.g., Pickering & Branigan, 1998). This lack of lexical boost in redundancy alignment may be explained due to the lexical overlap in Experiment 1 consisting of non-head lexical items (i.e., the adjectives). Syntactic alignment has been found to not be affected by the repetition of non-head lexical items (Carminati et al., 2019; Cleland & Pickering, 2003), and redundancy alignment may operate in a similar fashion. Note though that all three experiments in the study included lexical overlap of the head lexical items (i.e., the nouns). Therefore, if a lexical boost effect involving non-head adjectives does exist, then this effect may have been overridden by a lexical boost effect involving head nouns (as lexical boost effects involving heads are likely stronger than non-head lexical boost effects). Future research should compare redundancy alignment with and without overlapping head nouns and non-head adjectives to further examine if lexical boost effects play a role in this type of alignment.

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

Redundancy alignment could also be the outcome of a goal-oriented process, in which an individual imitates their interlocutor's linguistic (redundancy) usage in order to boost communicative success. This is consistent with communication accommodation theory, whereby speakers converge on the redundancy usage patterns of their interlocutor in order to meet the expectation of an interaction (Giles, 2008). In scenarios where the interlocutor uses redundancies (e.g., in color or size), the individual might deduce that the interlocutor prefers to use size/color (despite being redundant) to single out a target object and therefore tend to use redundancies in subsequent interactions in order for the interlocutor to better identify the target object. However, this goal-oriented alignment does not appear to depend on the interlocutor's identity (at least in our setup), as redundancy alignment towards computers was not greater than towards humans (with this patterning of results being present with participants who believed that their partner was a human/computer when told they would be conducting the task with a human/computer, respectively). This concurs with earlier research demonstrating that people tend to align syntactically at the same rate with computers and humans (Brennan, 1991; Cowan et al., 2015; Heyselaar et al., 2017), while contrasting with research showing that people tend to lexically align more with computers than humans (Branigan et al., 2004; Branigan et al., 2011; Bergmann, Branigan, & Kopp, 2015; Shen & Wang, 2023). Overall, this suggests that lexical alignment occurs both from an automatic mechanism such as priming and from a goal-directed mechanism such as interlocutor modeling, but alignment in syntax and redundancy usage seems to mainly come from an automatic mechanism.

Nonetheless, there was a tendency (albeit rather small) for people to use more redundancies (i.e., an overall increase in redundancy usage) with computers compared to humans. These results may arise from a goal-oriented effect of interlocutor modeling on general redundancy production rather than on redundancy alignment specifically. Here, people may use more redundancies towards computers as a strategy for increasing comprehension with a linguistically less competent interlocutor (i.e., the computer). This is consistent with the notion of people altering their utterances to a greater extent towards computers as a result of a goal-oriented process (e.g., Branigan et al., 2004). In addition, higher perceived competency of the interlocutor was associated with less redundancy usage, further supporting the notion that people may use more redundancies with interlocutors they perceive as less competent in order to aid them in comprehension (e.g., Deutsch & Pechmann, 1982; Rubio-Fernández, 2016; Arts et al., 2011). Furthermore, there was a difference in the perceived competency between humans and computers (Experiment 3), suggesting that the increased redundancy usage towards computer interlocutors was due to people trying to aid a linguistically less competent interlocutor.

Overall, these results of redundancy alignment not being modulated by interlocutor identity (i.e., computer vs. human) and general redundancy usage increasing slightly for computer interlocutors concur with past research showing both similarities and differences between HCI and HHI. In general, people seem to apply similar behavioral schemas in HCI as in HHI due to people viewing artificial entities as social actors (Krämer, 2005; Bartneck et al., 2009; Groom et al., 2011), approximating human-like competencies (McKee et al. 2023), having moral

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

standing (Malle et al., 2015, 2019), and having highly competent perceptual and linguistic abilities (Saryazdi et al., 2021), leading to people treating computers as humans (e.g., Nass et al., 1997). This is especially the case with advanced technologies such as large language models that approximate human competencies in areas such as language to a much larger extent than in the past (e.g., Barattieri di San Pietro et al., 2023). Likewise, our results generally show remarkable similarities between HCI and HHI, pointing to this general trend of these two interaction modalities converging.

However, small differences in redundancy usage between HCI and HHI also suggest that people still perceive differences between these interlocutors, with these perceptions affecting linguistic production. These perception differences were found in the post-experiment ratings, as well as in the experimental picture naming and matching task that is relatively easy to complete. Therefore, we propose that artificial intelligence artefacts (such as, large language models, robots) are treated as quasi-humans, due to the tendency for people to perceive and treat artificial entities as qualitatively human in the broad majority of circumstances but, in some cases, with subtle differences (i.e., small quantitative differences in a certain behavior), indicating that these entities are viewed as quasi-human but not fully human. Of course, linguistic alignment is only one small area in the broad realm of interactional possibilities, and future research should set out to examine the extent to which people interact with artificial entities as quasi-human. Nonetheless, these results concur with past findings showing that people mindlessly apply social norms onto a-social entities such as computers and hence treat computers as human (Nass & Moon, 2000, Lee, 2010) and that people treat artificial entities as humans when these entities perform a task as expected (Lee, 2024).

Additional modulators of redundancy usage and alignment

Upon comparing alignment rates towards interlocutors who either gave written or spoken responses, alignment towards spoken redundant interlocutors was stronger compared to written redundant interlocutors. This effect might be attributed to redundancies being more noticeable in spoken form. When redundancies are heard in speech, people may be more aware of this linguistic behavior and thus are more likely to include redundancies in their subsequent utterances (even when using a different response medium than their interlocutor, in this case typing). This stronger alignment may also be a consequence of increased attention directed towards the identity of the interlocutor. Hearing computer or human speech may draw more attention to the fact that the interlocutor is a computer or a human, respectively, and could therefore accentuate the alignment differences between these two types of interlocutors. However, past research has found that alignment rates do not change when hearing human-like versus more artificial-sounding voices (Cowan & Branigan, 2015), suggesting that hearing a voice might not necessarily lead to more attention being paid to interlocutor-specific characteristics.

Redundancy alignment may also be modulated by individual differences in addition to the identity of one's interlocutor. Higher ToM was associated with increased redundancy alignment, which suggests that ToM may play a role in redundancy alignment whereby superior ToM abilities aid the understanding of the

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

mental states of others (Premack & Woodruff, 1978) and therefore facilitate a better discernment of an interlocutor's redundancy usage, creating an opportunity to align with this linguistic behavior. However, as ToM was not experimentally manipulated, the direction of causation between ToM and redundancy alignment cannot be established, and another likely possibility is that these are linked via a third causative factor such as attention. In addition, ToM does not seem to play a role in establishing a more definite interlocutor model, as ToM did not modulate the rates in which speakers aligned to the redundancy usage patterns of computers compared to humans. This may be due to the fact that, even though people are more knowledgeable in general about humans compared to computers (Epley et al., 2007), discerning the differences between humans and computers probably is relatively easy, thereby making any increase in ToM skills not beneficial in creating adequate mental models of these two types of interlocutors. The measure of ToM as well was found to not have high internal reliability, further suggesting that the results involving ToM should be interpreted with caution.

The varying competency perceptions that speakers had of their interlocutors also affect alignment, whereby speakers aligned more towards non-redundant interlocutors who they perceived as more competent, with this trend absent towards redundant interlocutors. This effect is driven by speakers using fewer redundancies overall with non-redundant interlocutors who were perceived as more competent, suggesting that redundancy usage is being used as a strategy to increase communicative success with less competent interlocutors. These linguistic behaviors are consistent with the tentative finding in this study that speakers in general use more redundancies with computers than humans, possibly as a strategy to aid computers (i.e., the perceived less competent interlocutor) in comprehension compared to humans (i.e., the perceived more competent interlocutor). However, this finding of increased redundancy usage towards computers was only significant in Experiment 1, marginally significant in Experiment 2, and non-significant in Experiment 3, suggesting that while in some cases people may use more redundancies towards computers, redundancy usage is broadly similar towards these interlocutors. Therefore, it seems that redundancy usage is driven to a larger extent by competency perceptions that individuals have of the comprehender and to a lesser extent by competency perception differences across humans and computers.

While throughout this study we have differentiated computer and human interlocutors mainly through competence, it should be noted that these interlocutors differ on a myriad of traits. Competency was examined primarily in this study, as much past research comparing HCI and HHI has looked at competency differences to explain perceptual and behavioral differences between these interaction modalities (e.g., Branigan et al., 2004; Branigan et al., 2011; Pearson et al., 2006) and found that competency is an influential predictor when comparing HCI to HHI. However, other differences (e.g., agency, sociability, etc.) may modulate differences between HCI and HHI, including in redundancy usage and alignment. While looking at traits beyond competency is beyond the scope of this study, future research should examine how a broader range of traits cause divergences between HCI and HHI.

In addition, as the interactions progressed, people tended to decrease their use of redundancies with both redundant and non-redundant interlocutors (all

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

experiments), which suggests that redundancy usage is effortful (at least in typing, which was the response medium across all experiments). This intuitively seems plausible, as the usage of redundancies involves noticing additional properties of a referent and producing a larger volume of linguistic output. Note that the current study did not have participants verbally respond; however, verbal responses are still predicted to show the same time course pattern as written responses. While verbal responses may be somewhat less effortful to produce than written responses, producing verbal redundancies still involves increased attentional and motor effort in formulating and producing longer utterances and therefore is likely to show a similar (but perhaps slower) decrease in redundancy usage over time. However, this finding contrasts with the observation that redundancy alignment intensifies as the interaction continues (Loy & Smith, 2021). Nonetheless, the Loy and Smith study did not measure redundancy in the intervening filler trials where individuals had the opportunity to produce redundancies, making the true rate of redundancy alignment across all utterances during the interactions unclear.

While it is plausible that the effort required to produce redundancies causes redundancy usage to decrease over time, this study did not explicitly test this prediction, and other factors may have caused this decrease. As the interaction progressed in the experimental trials, participants may have developed the belief that redundancy usage may no longer be helpful for their interlocutor, as the trials were programmed to go smoothly and without concerns or signs of comprehension difficulty from the interlocutor. Therefore, this belief in task competence of the interlocutor over time may have driven the reduction in redundancy usage over time and not from redundancies being effortful. Likewise, it is possible that both of these factors are at play in reducing redundancy usage over time, and future research should conduct experiments to test these predictions.

The rate of redundancy usage reduced more slowly with redundant interlocutors, suggesting that people make an attempt to align with redundant interlocutors, but this alignment still decreases over time. In contrast, the reduction in redundancies over time when interacting with non-redundant interlocutors may merely reflect a process where people produce more efficient (i.e., less effortful) utterances as the interaction progresses, or due to people believing that redundancies are no longer helpful for the interlocutor, rather than a result of redundancy alignment. This aligns with the inverse frequency effect, where people align more strongly with less frequent than with more frequent constructions (e.g., Hartsuiker & Westenberg, 2000; Scheepers, 2003). Various studies (Deutsch & Pechmann, 1982; Engelhardt et al., 2006), along with the present one, have found redundant utterances to be less frequent than non-redundant utterances. Therefore, alignment towards redundant utterances may be stronger than towards non-redundant ones. The slower reduction of redundant utterances when interacting with redundant interlocutors than non-redundant ones might thus reflect stronger alignment towards redundancies. In such cases, this increased alignment could mitigate some of the effects of a general decrease in redundancy usage over time.

The overall reduction in redundancies and the absence of increased alignment over time are in line with the concept of rapid decay of residual activation (Pickering & Branigan, 1998; Branigan, Pickering, & Cleland, 1999). Given this rapid decay, redundancy alignment does not appear to accumulate in strength even as the

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

interaction progresses. Instead, the alignment is primarily driven by the direct recent exposure to redundant or non-redundant utterances in the matching trials.

Applications

As well as informing linguistic theory, the study findings also can inform designers of computer natural language systems on the nature of linguistic interactions (e.g., Stoyanchev & Stent, 2009). As rates of redundancy usage seem to be dropping across time, these systems can be engineered to expect this linguistic behavior, which can therefore aid in system comprehension and task performance by informing on the linguistic units that are likely to be produced by the user. More generally, the finding of linguistic alignment in redundancy usage creates the possibility for systems to subtly direct users to produce or not produce redundancy, which can constrain the potential linguistic space and aid in comprehension for speech recognition systems (Cowan et al., 2015). Thus, having users use or avoid redundancy due to linguistic alignment, and concurrently systems with expectations that these redundancy patterns will be used, can greatly aid speech recognition and increase communicative success. In addition, when a system detects increased redundancy usage directed towards it from a user, this may be an indication that the user does not think that the system has comprehended correctly or that the system is not achieving the communicative goal. This can be inferred from the finding that greater redundancy usage was found when people perceived their interlocutor as having lower competence. Therefore, in these situations systems can choose a different (but also likely) comprehension of the linguistic input received and/or provide a different response to increase the likelihood of communicative success.

Conclusion

In sum, this study shows that speakers frequently produce redundancies, and they align their redundancy usage with that of their conversation partner. The degree of redundancy alignment is modulated by several factors, including the perceived competence of the conversation partner, the modality of the utterances produced by the interlocutor (spoken versus written), individual differences in ToM, and the time course of the interaction. The alignment of redundancies (and redundancy usage in general) seems to originate from both an automatic, interlocutor-independent process and a more goal-oriented process, in which individuals adjust their language use based on the perceived linguistic competence of their interlocutor. In doing so, they may use redundancies as a strategy to facilitate successful communication. Overall, people broadly treat computers the same as humans, suggesting that people view computers and similar technologies as quasi-human. (i.e., almost but not quite human). Future research should investigate these various factors in more depth, as well as exploring the potential interplay between different levels of linguistic behavior, such as lexical and syntactic alignment, in order to gain a comprehensive understanding of how speakers adjust their language in real-time conversation.

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

Replication package. The data and analytical scripts can be found on Open Science Framework (https://osf.io/2kzn9/).

Acknowledgements. This work was supported by a General Research Fund grant (Project Number: 14600220).

Competing interests. We have no declarations of interest.

References

Arts, A., Maes, A., Noordman, L., & Jansen, C. (2011). Overspecification in written instruction. *Linguistics*, **49**(3), 555–574. https://doi.org/10.1515/ling.2011.017

Bannon, J., Saryazdi, R., & Chambers, C. G. (2020). Designing referential descriptions for children, young adults, and computers: a comprehensive examination of talker informativity. In *Proceedings of the 42nd Annual Conference of the Cognitive Science Society*. Cognitive Science Society.

Barattieri di San Pietro, C., Frau, F., Mangiaterra, V., & Bambini, V. (2023). The pragmatic profile of ChatGPT: assessing the communicative skills of a conversational agent. *Sistemi Intelligenti*, **35**(2), 379–400. https://doi.org/10.1422/108136

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, **42**(2), 241–251. https://doi.org/10.1515/ling.2011.017

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, **1**, 71–81. https://doi.org/10.1007/s12369-008-0001-3

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*. https://doi.org/10.48550/arXiv.1506.04967

Bavelas, J. B., Black, A., Lemery, C. R., & Mullett, J. (1986). "I show how you feel": motor mimicry as a communicative act. *Journal of Personality and Social Psychology*, **50**(2), 322–329. https://doi.org/10.1037/0022-3514.50.2.322

Bell, L., Gustafson, J., & Heldner, M. (2003). Prosodic adaptation in human-computer interaction. *Proceedings of ICPHS*, **3**, 833–836.

Bergmann, K., Branigan, H. P., & Kopp, S. (2015). Exploring the alignment space–lexical and gestural alignment with real and virtual humans. *Frontiers in ICT*, **2**, 7. https://doi.org/10.3389/fict.2015.00007

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (1999). Syntactic priming in written production: evidence for rapid decay. *Psychonomic Bulletin & Review*, **6**, 635–640. https://doi.org/10.3758/BF03212972

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, **75**(2), B13–B25. https://doi.org/10.1016/S0010-0277(99)00081-5

Branigan, H. P., Pickering, M. J., Pearson, J., McLean, J. F., & Brown, A. (2011). The role of beliefs in lexical alignment: evidence from dialogs with humans and computers. *Cognition*, **121**(1), 41–57. https://doi.org/10.1016/j.cognition.2011.05.011

Branigan, H. P., Pickering, M. J., Pearson, J., McLean, J. F., & Nass, C. I. (2003). Syntactic alignment between computers and people: the role of belief about mental states. *Proceedings of the 25th Annual Conference of the Cognitive Science Society*, **31**, 186–191.

Branigan, H. P., Pickering, M. J., Pearson, J., McLean, J. F., Nass, C. I., & Hu, J. (2004). Beliefs about mental states in lexical and syntactic alignment: evidence from human-computer dialogs. In *Proceedings of the CUNY Conference on Human Sentence Processing*. College Park, MD: University of Maryland.

Brennan, S. E. (1991). Conversation with and through computers. *User Modeling and User-Adapted Interaction*, **1**, 67–86. https://doi.org/10.1007/BF00158952

Brennan, S. E. (1996). Lexical entrainment in spontaneous dialog. *Proceedings of ISSD*, **96**, 41–44.

Cai, Z. G., Sun, Z., & Zhao, N. (2021). Interlocutor modelling in lexical alignment: the role of linguistic competence. *Journal of Memory and Language*, **121**, 104278. https://doi.org/10.1016/j.jml.2021.104278

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

Carminati, M. N., van Gompel, R. P., & Wakeford, L. J. (2019). An investigation into the lexical boost with nonhead nouns. *Journal of Memory and Language*, **108**, 104031.

Cirillo, G., Runnqvist, E., Strijkers, K., Nguyen, N., & Baus, C. (2022). Conceptual alignment in a joint picture-naming task performed with a social robot. *Cognition*, **227**, 105213. https://doi.org/10.1016/j.cognition.2022.105213

Cleland, A. A., & Pickering, M. J. (2003). 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), 214–230. https://doi.org/10.1016/S0749-596X(03)00060-3

Cohn, M., & Zellou, G. (2021). Prosodic differences in human-and Alexa-directed speech, but similar local intelligibility adjustments. *Frontiers in Communication*, **6**, 675704. https://doi.org/10.3389/fcomm.2021.675704

Cowan, B. R., & Branigan, H. P. (2015). Does voice anthropomorphism affect lexical alignment in speech-based human-computer dialogue?. *Proceedings of the Annual Conference of the International Speech Communication Association (Interspeech)*, **2015**, 155–159.

Cowan, B. R., Branigan, H. P., Obregón, M., Bugis, E., & Beale, R. (2015). Voice anthropomorphism, interlocutor modelling and alignment effects on syntactic choices in human– computer dialogue. *International Journal of Human-Computer Studies*, **83**, 27–42. https://doi.org/10.1016/j.ijhcs.2015.05.008

Dehn, D. M., & Van Mulken, S. (2000). The impact of animated interface agents: a review of empirical research. *International Journal of Human-Computer Studies*, **52**(1), 1–22. https://doi.org/10.1006/ijhc.1999.0325

Deutsch, W., & Pechmann, T. (1982). Social interaction and the development of definite descriptions. *Cognition*, **11**(2), 159–184. https://doi.org/10.1016/0010-0277(82)90024-5

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, **11**(1), 86–89. https://doi.org/10.1111/1467-9280.00221

Engelhardt, P. E., Bailey, K. G., & Ferreira, F. (2006). Do speakers and listeners observe the Gricean Maxim of quantity?. *Journal of Memory and Language*, **54**(4), 554–573. https://doi.org/10.1016/j.jml.2005.12.009

Engelhardt, P. E., Demiral, Ş. B., & Ferreira, F. (2011). Over-Specified referring expressions impair comprehension: An ERP study. *Brain and Cognition*, **77**(2), 304–314. https://doi.org/10.1016/j.bandc.2011.07.004

Engelhardt, P. E., & Ferreira, F. (2014). Do speakers articulate over-described modifiers differently from modifiers that are required by context? Implications for models of reference production. *Language, Cognition and Neuroscience*, **29**(8), 975–985. https://doi.org/10.1080/01690965.2013.853816

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: a three-factor theory of anthropomorphism. *Psychological Review*, **114**(4), 864. https://doi.org/10.1037/0033-295X.114.4.864

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, **39**(2), 175–191. https://doi.org/10.3758/BF03193146

Fukumura, K., & Carminati, M. N. (2022). Overspecification and incremental referential processing: an eye-tracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **48**(5), 680. https://doi.org/10.1037/xlm0001015

Fukumura, K., & van Gompel, R. P. (2017). How do violations of Gricean maxims affect reading?. *Journal of Memory and Language*, **95**, 1–18. https://doi.org/10.1016/j.jml.2017.01.008

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: a study in conceptual and semantic co-ordination. *Cognition*, **27**(2), 181–218. https://doi.org/10.1016/0010-0277(87)90018-7

Gessinger, I., Möbius, B., Andreeva, B., Raveh, E., & Steiner, I. (2019). Phonetic accommodation in a Wizard-of-Oz experiment: intonation and segments. *Proceedings of the Annual Conference of the International Speech Communication Association (Interspeech)*, **2019**, 301–305.

Giles, H. (1973). Accent mobility: a model and some data. *Anthropological Linguistics*, **15**(2), 87–105.

Giles, H. (2008). Accommodating translational research. *Journal of Applied Communication Research*, **36**(2), 121–127. https://doi.org/10.1080/00909880801922870

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: communication, context, and consequence. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of accommodation: Developments in applied sociolinguistics* (pp. 1–68). Cambridge: Cambridge University Press.

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

Giles, H., Edwards, A. L., & Walther, J. B. (2023). Communication accommodation theory: past accomplishments, current trends, and future prospects. *Language Sciences*, **99**, 101571. https://doi.org/10.1016/j.langsci.2023.101571

Goldin-Meadow, S., & Alibali, M. W. (2013). Gesture's role in speaking, learning, and creating language. *Annual Review of Psychology*, **64**, 257–283. https://doi.org/10.1146/annurev-psych-113011-143802

Goudbeek, M., & Krahmer, E. (2012). Alignment in interactive reference production: content planning, modifier ordering, and referential overspecification. *Topics in Cognitive Science*, **4**(2), 269–289. https://doi.org/10.1111/j.1756-8765.2012.01186.x

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Speech acts* (pp. 41–58). Leiden: Brill.

Groom, V., Srinivasan, V., Bethel, C. L., Murphy, R., Dole, L., & Nass, C. (2011, May). Responses to robot social roles and social role framing. In *2011 International Conference on Collaboration Technologies and Systems (CTS)* (pp. 194–203). https://doi.org/10.1109/CTS.2011.5928687

Hartsuiker, R. J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. *Cognition*, **75**(2), B27–B39. https://doi.org/10.1016/S0010-0277(99)00080-3

Heyselaar, E., Hagoort, P., & Segaert, K. (2017). In dialogue with an avatar, language behavior is identical to dialogue with a human partner. *Behavior Research Methods*, **49**, 46–60. https://doi.org/10.3758/s13428-015-0688-7

Kopp, S., Gesellensetter, L., Krämer, N. C., & Wachsmuth, I. (2005). A conversational agent as museum guide: design and evaluation of a real-world application. In T. Panayiotopoulos, J. Gratch, R. Aylett, D. Ballin, & P. Olivier (Eds.), *Intelligent virtual agents: 5th International Working Conference, IVA 2005, Kos, Greece, September 12–14, 2005. Proceedings* (pp. 329–343). Berlin: Springer Berlin Heidelberg.

Krämer, N. C. (2005). Social communicative effects of a virtual program guide. In T. Panayiotopoulos, J. Gratch, R. Aylett, D. Ballin, & P. Olivier (Eds.), *Intelligent virtual agents: 5th International Working Conference, IVA 2005, Kos, Greece, September 12–14, 2005. Proceedings* (pp. 442–453). Berlin: Springer Berlin Heidelberg.

Lee, E. J. (2010). What triggers social responses to flattering computers? Experimental tests of anthropomorphism and mindlessness explanations. *Communication Research*, **37**(2), 191–214. https://doi.org/10.1177/0093650209356389

Lee, E. J. (2024). Minding the source: toward an integrative theory of human–machine communication. *Human Communication Research*, **50**(2), 184–193. https://doi.org/10.1093/hcr/hqad034

Levinson, S. C. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press.

Loy, J. E., & Smith, K. (2021). Speakers align with their partner's overspecification during interaction. *Cognitive Science*, **45**(12), e13065. https://doi.org/10.1111/cogs.13065

Malle, B. F., Magar, S. T., & Scheutz, M. (2019). AI in the sky: how people morally evaluate human and machine decisions in a lethal strike dilemma. In M. I. A. Ferreira, J. S. Sequeira, G. S. Virk, M. O. Tokhi, & E. E. Kadar (Eds.), *Robotics and Well-Being* (pp. 111–133). Cham: Springer Verlag.

Malle, B. F., Scheutz, M., Arnold, T., Voiklis, J., & Cusimano, C. (2015, March). Sacrifice one for the good of many? People apply different moral norms to human and robot agents. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction* (pp. 117–124). https://doi.org/10.1145/2696454.269645

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing type I error and power in linear mixed models. *Journal of Memory and Language*, **94**, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

McKee, K. R., Bai, X., & Fiske, S. T. (2023). Humans perceive warmth and competence in artificial intelligence. *Iscience*, **26**(8), 107256.

Mengesha, Z., Heldreth, C., Lahav, M., Sublewski, J., & Tuennerman, E. (2021). "I don't think these devices are very culturally sensitive."—Impact of automated speech recognition errors on African Americans. *Frontiers in Artificial Intelligence*, **4**, 725911. https://doi.org/10.3389/frai.2021.725911

Nass, C., & Moon, Y. (2000). Machines and mindlessness: social responses to computers. *Journal of Social Issues*, **56**(1), 81–103. https://doi.org/10.1111/0022-4537.00153

Nass, C. I., Moon, Y., & Green, N. (1997). Are machines gender neutral? Gender-stereotypic responses to computers with voices. *Journal of Applied Social Psychology*, **27**(10), 864–876. https://doi.org/10.1111/j.1559-1816.1997.tb00275.x

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

Oviatt, S. (1995). Predicting spoken disfluencies during human-computer interaction. *Computer Speech and Language*, **9**(1), 19–36. https://doi.org/10.1006/csla.1995.0002

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America*, **119**(4), 2382–2393. https://doi.org/10.1121/1.2178720

Parise, S., Kiesler, S., Sproull, L., & Waters, K. (1999). Cooperating with life-like interface agents. *Computers in Human Behavior*, **15**(2), 123–142. https://doi.org/10.1016/S0747-5632(98)00035-1

Pearson, J., Hu, J., Branigan, H. P., Pickering, M. J., & Nass, C. I. (2006, April). Adaptive language behavior in HCI: how expectations and beliefs about a system affect users' word choice. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1177–1180). https://doi.org/10.1145/1124772.1124948

Pechmann, T. (1989). Incremental speech production and referential overspecification. *Linguistics*, **27**(1), 89–110. https://doi.org/10.1515/ling.1989.27.1.89

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: evidence from syntactic priming in language production. *Journal of Memory and Language*, **39**(4), 633–651. https://doi.org/10.1006/jmla.1998.2592

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, **27**(2), 169–190. https://doi.org/10.1017/S0140525X04000056

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. *Behavioral and Brain Sciences*, **1**(4), 515–526. https://doi.org/10.1017/S0140525X00076512

Rao, X., Wu, H., & Cai, Z. G. (2024). Comprehending semantic and syntactic anomalies in LLM- versus human-generated texts: an ERP study. *OSF Preprint*, 8 October. https://doi.org/10.31234/osf.io/kvbg2

Rubio-Fernández, P. (2016). How redundant are redundant color adjectives? An efficiency-based analysis of color overspecification. *Frontiers in Psychology*, **7**, 153. https://doi.org/10.3389/fpsyg.2016.00153

Rubio-Fernández, P. (2021). Color discriminability makes over-specification efficient: theoretical analysis and empirical evidence. *Humanities and Social Sciences Communications*, **8**(1), 1–15. https://doi.org/10.1057/s41599-021-00818-6

Saryazdi, R., Nuque, J., & Chambers, C. (2021). Younger and older speakers' use of linguistic redundancy with a social robot. *Proceedings of the Annual Meeting of the Cognitive Science Society*, **43**(43), 1823–1829. https://escholarship.org/uc/item/9zt616fw

Saryazdi, R., Nuque, J., & Chambers, C. G. (2022). Linguistic redundancy and its effects on younger and older adults' real-time comprehension and memory. *Cognitive Science*, **46**(4), e13123. https://doi.org/10.1111/cogs.13123

Scheepers, C. (2003). Syntactic priming of relative clause attachments: persistence of structural configuration in sentence production. *Cognition*, **89**(3), 179–205. https://doi.org/10.1016/S0010-0277(03)00119-7

Sedivy, J. C. (2003). Pragmatic versus form-based accounts of referential contrast: evidence for effects of informativity expectations. *Journal of Psycholinguistic Research*, **32**, 3–23. https://doi.org/10.1023/A:1021928914454

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, **71**(2), 109–147. https://doi.org/10.1016/S0010-0277(99)00025-6

Shechtman, N., & Horowitz, L. M. (2003, April). Media inequality in conversation: how people behave differently when interacting with computers and people. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 281–288). https://doi.org/10.1145/642611.642661

Shen, H., & Wang, M. (2023). Effects of social skills on lexical alignment in human-human interaction and human-computer interaction. *Computers in Human Behavior*, **143**, 107718. https://doi.org/10.1016/j.chb.2023.107718

Sproull, L., Subramani, M., Kiesler, S., Walker, J. H., & Waters, K. (1996). When the interface is a face. *Human-Computer Interaction*, **11**(2), 97–124. https://doi.org/10.1207/s15327051hci1102_1

Stent, A. J., Huffman, M. K., & Brennan, S. E. (2008). Adapting speaking after evidence of misrecognition: Local and global hyperarticulation. *Speech Communication*, **50**(3), 163–178. https://doi.org/10.1016/j.specom.2007.07.005

Stoyanchev, S., & Stent, A. (2009, June). Lexical and syntactic adaptation and their impact in deployed spoken dialog systems. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of*

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

the North American Chapter of the Association for Computational Linguistics, Companion Volume: Short Papers (pp. 189–192).

Suffill, E., Kutasi, T., Pickering, M. J., & Branigan, H. P. (2021). Lexical alignment is affected by addressee but not speaker nativeness. *Bilingualism: Language and Cognition*, **24**(4), 746–757. https://doi.org/10.1017/S1366728921000092

Tia, B., Saimpont, A., Paizis, C., Mourey, F., Fadiga, L., & Pozzo, T. (2011). Does observation of postural imbalance induce a postural reaction?. *PloS One*, **6**(3), e17799. https://doi.org/10.1371/annotation/aff2c1f3-08af-4a36-bf53-0df2bfd9c320

Tooby, J., & Cosmides, L. (1995). Foreword. In S. Baron-Cohen (Ed.), *Mindblindness: An Essay on Autism and Theory of Mind* (pp. 11–18). Cambridge, MA: MIT Press.

Van Gompel, R. P., Van Deemter, K., Gatt, A., Snoeren, R., & Krahmer, E. J. (2019). Conceptualization in reference production: probabilistic modeling and experimental testing. *Psychological Review*, **126**(3), 345. https://doi.org/10.1037/rev0000138

Van Lierop, K., Goudbeek, M., & Krahmer, E. (2012). Conceptual alignment in reference with artificial and human dialogue partners. *Proceedings of the Annual Meeting of the Cognitive Science Society*, **34**(34), 1066–1071.

Wahn, B., Berio, L., Weiß, M., & Newen, A. (2023). Try to see it my way: humans take the level-1 visual perspective of humanoid robot avatars. *International Journal of Social Robotics*, **17**(3), 523–534. https://doi.org/10.1007/s12369-023-01036-7

Wu, H., & Cai, Z. G. (2024). Speaker effects in spoken language comprehension. *arXiv preprint arXiv:2412.07238*. https://doi.org/10.48550/arXiv.2412.07238

Zhao, X., & Malle, B. F. (2022). Spontaneous perspective taking toward robots: the unique impact of humanlike appearance. *Cognition*, **224**, 105076. https://doi.org/10.1016/j.cognition.2022.105076

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

Appendix

Figures

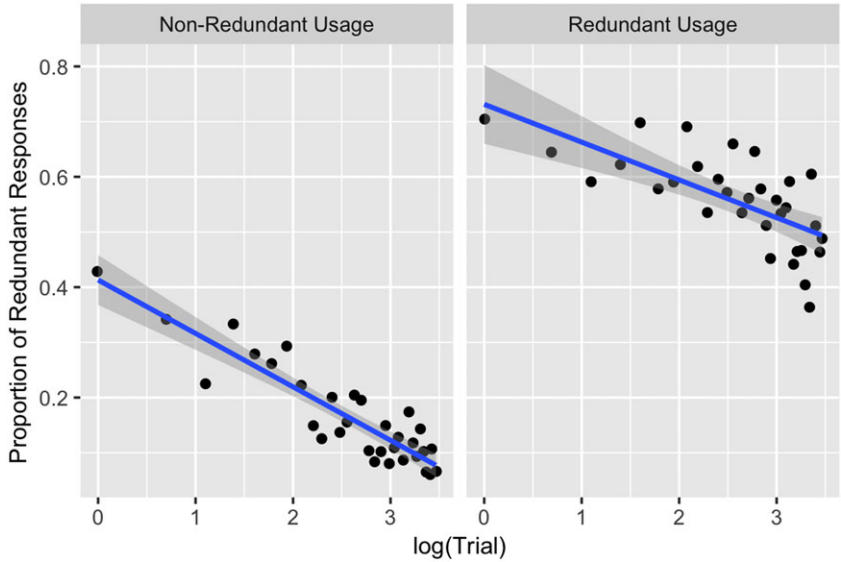

Figure A. Proportion of redundant responses for usage and trial for Experiment 1 (shading shows the 95% CI).

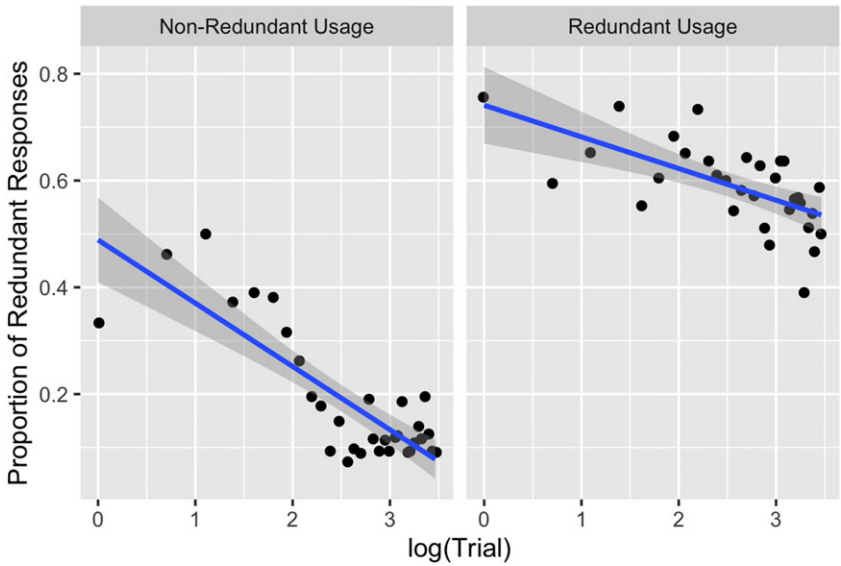

Figure B. Proportion of redundant responses for usage and trial for Experiment 2 (shading shows the 95% CIs).

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press

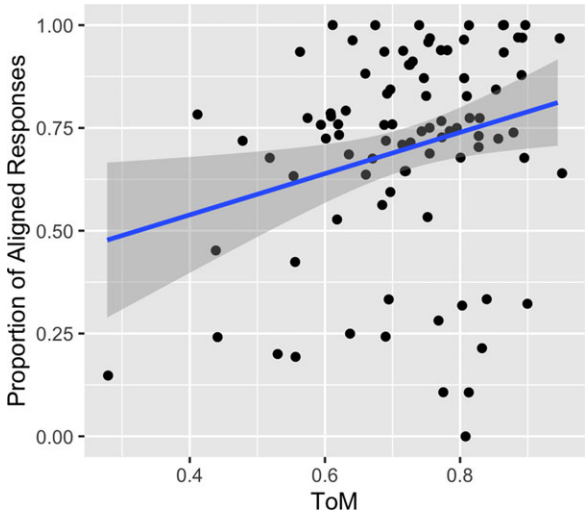

Figure C. Proportion of aligned responses for ToM for Experiment 2 (shading shows the 95% CIs).

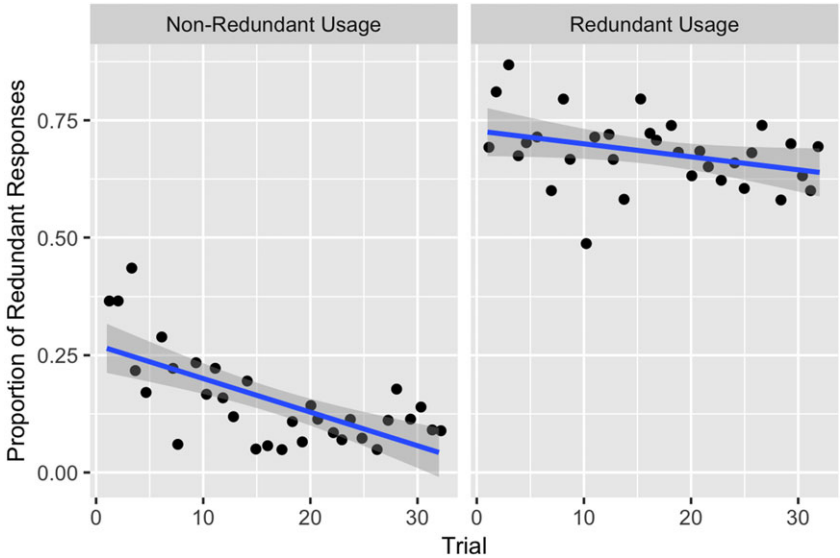

Figure D. Proportion of redundant responses for usage and trial for Experiment 3 (shading shows the 95% CIs).

Cite this article: Dunn, M. S. & Cai, Z. G. (2025). Linguistic alignment of redundancy usage in human-human and human-computer interaction. *Applied Psycholinguistics*. https://doi.org/10.1017/S0142716425100118

https://doi.org/10.1017/S0142716425100118 Published online by Cambridge University Press