Title
Shared context and lexical alignment: an experimental investigation

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Mudd, Katie
Schouwstra, Marieke

Publication Date
2024

Peer reviewed

Shared context and lexical alignment: an experimental investigation

Katie Mudd (katie.mudd@ai.vub.ac.be)
Institute for Logic, Language and Computation, University of Amsterdam
Amsterdam, the Netherlands

Marieke Schouwstra (m.schouwstra@uva.nl)
Institute for Logic, Language and Computation, University of Amsterdam
Amsterdam, the Netherlands

Abstract

What drives lexical alignment in the context of language emergence? We test the theory that limited context promotes alignment, because individuals cannot make use of iconic mappings between shared meanings and forms. Using a novel referential communication paradigm where participants use pre-recorded gesture videos to communicate, we test different context conditions. We find, unexpectedly, no alignment differences between dyads with shared context and dyads with limited context, even though the former have fewer communicative errors. Importantly, we do observe differences when it comes to the iconic strategies used: less shared context promotes the use of (shared) visual iconicity.

Keywords: shared context; alignment; lexical variation; iconicity; experimental semiotics; language evolution

Introduction

Languages are governed by conventions (norms or principles governing linguistic structure) but they also show variation, and in the study of linguistic structure, conventions have traditionally received more attention than variation. Studying language emergence, research in experimental semiotics has focused on how linguistic systems go from a high degree of unstructured variation (where individuals use different words to refer to a concept) to alignment (where individuals all use the same word to refer to an object). For example, in an experiment by Fay, Garrod, Roberts, and Swoboda (2010), micro-societies who communicated about a set of concepts via drawings tended to align on drawings over time, with each micro-society having unique drawing norms. But which cultural variants get selected over others? This selection has been proposed to be influenced by a variety of cognitive strategies (Boyd & Richerson, 1988; Mesoudi, 2011). One proposition is the *content-bias*, suggesting that participants are more likely to align on variants that are easier to use or learn (Boyd & Richerson, 1988; for a model supporting this see Tamariz et al., 2014), which may offer advantages in processing (Christiansen & Chater, 2008).

Besides the question of what cognitive strategies drive alignment, it is worth asking why individuals would want to align in the first place. It has long been suggested that there is a strong connection between alignment and communicative success (Garrod & Pickering, 2004). Supporting this theory, previous experiments, such as interactive map tasks (introduced by Brown et al., 1984) where individuals need to communicate about the location of objects on a map, have shown that alignment decreases communicative error (Fusaroli & Tylén, 2016). Intuitively, it is plausible that limitations in working memory affect alignment: more frequent variants may be the most easily accessible (Hudson Kam & Chang, 2009). As such, it appears that certain cognitive mechanisms favor alignment.

Here we are interested in the opposite question: are there cognitive mechanisms that deter communication systems from aligning while still allowing individuals to achieve successful communication? One plausible cognitive capacity is the ability to make use of *iconicity*, which refers to the degree of similarity between the mapping of a form and its meaning (Perniss, Thompson, & Vigliocco, 2010; Taub, 2001). Research in experimental semiotics in the manual and vocal modalities has emphasized the role of iconicity in bootstrapping communication systems (Fay, Arbib, & Garrod, 2013; Fay, Ellison, & Garrod, 2014; Macuch Silva, Holler, Ozyurek, & Roberts, 2020). Research from signed and spoken languages also supports this (Perniss et al., 2010). Emmorey (2014) suggests that the previous longstanding hypothesis that iconicity decays in favor of arbitrariness (e.g., Aronoff & Sandler, 2005) is likely to be incorrect, given the critical role that iconicity plays in linguistic systems even after they have existed for a long period of time. Crucially, iconicity is subjective and an individual-level phenomenon, relying on the sensory-motor, perceptual, cultural and linguistic experiences of signers (Occhino, Anible, Wilkinson, & Morford, 2017). We are interested in how the subjective nature of iconicity interacts with alignment, drawing inspiration from observations about variation in sign languages and methods from experimental semiotics.

Given the subjective nature of iconicity (Occhino et al., 2017), it is important to take into account the amount of *shared context* interlocutors have. Shared context has been defined as the common amount of relevant information that interlocutors have (Winters, Kirby, & Smith, 2018; Müller, Winters, & Morin, 2019).[1] Information that is shared can range from visual information (Müller et al., 2019) to general cognitive information, and specific information that is culturally shared. Literature from sign language linguistics has pointed out interesting links between shared context and

[1]There are subtle differences between shared context and related terms such as *common ground* (Clark, 1996), *mutual cognitive environment* (Sperber & Wilson, 1996).

iconicity, and how they affect lexical alignment.

Certain sign languages appear to show less lexical alignment than others, and there have been several propositions to explain this (Meir, Israel, Sandler, Padden, & Aronoff, 2012). For example, in small, insular communities individuals may be able to keep track of each other's idiosyncratic expressions and therefore tolerate a high degree of lexical variation (de Vos, 2011; Thompson, Raviv, & Kirby, 2020), but recent cross-linguistic comparisons find no direct relationship between population size and lexical variation (Lutzenberger, Mudd, Stamp, & Schembri, 2023). A second explanation that we study here relates to shared context, iconicity and alignment, put forth by Tkachman and Hudson Kam (2020). At first glance, this explanation may seem counter-intuitive; a modeling effort by Mudd, de Vos, and de Boer (2022) provides a computational implementation, exploring the dynamics between parts of this theory. A summary of the model is visualised in Figure 1. In this model, it is assumed that lexical variation in a new linguistic system is high (with different individuals in the population improvising different forms). Iconicity in this model is represented as the overlap between the form and the meaning. As in classic language game models (e.g., Steels, 1995), interlocutors do not understand each other if they do not use the same forms and they have no other mechanisms to overcome this. Over time, repeated unsuccessful communication leads to alignment across the population. In the model by Mudd et al. (2022), this occurs in the limited context condition: the high degree of variation initially present is completely reduced in order for agents to successfully communicate. In contrast, in the shared context condition (where individuals have shared meanings), despite the high degree of variation present, interlocutors can retrieve meanings from forms (using iconicity) so they can successfully communicate and therefore do not need to align.

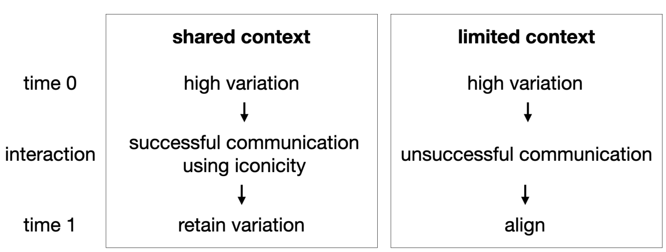

Figure 1: The relationship between lexical alignment and shared context (mediated by iconicity) as operationalized in Mudd et al. (2022).

The current study

The current study experimentally tests the relationship between shared context and lexical variation, as formalized in the agent-based model by Mudd et al. (2022). This model has shown that when a population has shared context, a high degree of lexical variation can be maintained. In doing this, it makes assumptions about the cognitive processes that individuals make when faced with lexical variation and different degrees of shared information. Here we wish to assess the validity of these cognitive assumptions by testing participants in a controlled setting, bridging research between observations from sign languages and models about the influence of social structure on lexical variation. Specifically, we focus here on the cognitive affordance of iconicity in mediating the relationship between shared context and lexical alignment.

To study iconicity in our experiment, we follow Ortega and Özyürek (2020) who outline different types of iconicity observed in gesture research: in the *acting* strategy, an individual shows how objects are manipulated with their body representing the actor (e.g., "smoking" represented with the action of moving one's fingers to the mouth as if holding a cigarette). In the *drawing* strategy, the outline of the object is traced (e.g., "house" represented by tracing the outline of a house). In the *representing* strategy, the hand represents the form of the referent (e.g., "mountain" represented by holding both of one's arms up to show the outline of the mountain). There is a conceptual connection between iconicity and shared context (Occhino et al., 2017); if I didn't know how people smoke, then the action gesture of moving your fingers near your lips would be meaningless to me. However, other strategies, like drawing, may be more easily guessed if we can both see the referent. As such, some types of iconicity require shared cultural knowledge while other types of iconicity may require only shared visual knowledge.

In our experiment, dyads take part in a referential communication game (first introduced by Krauss and Weinheimer, 1964), in which a director must communicate to a matcher about a target object. We alter this classic paradigm by having participants communicate about about unfamiliar objects using prerecorded gestures. These gestures can be iconic in two ways: some gestures are iconic in that the form of the gesture can be retrieved from visual properties of the object. All participants have access to this kind of iconicity. Secondly, gesture-object mappings can also be iconic depending on the 'cultural' information that a participant has about an object. Dyads who have shared context are trained on cultural knowledge that allows them to make the same mapping between objects and gestures as their partner. The crucial manipulation in our experiment is the amount of shared context that participants in each dyad have, which we test in three different conditions with varying levels of shared information.

We hypothesize that dyads who have a high degree of shared context will exhibit more lexical variation over time, as different iconic form-meaning mappings can be made thanks to their shared context. On the other hand, we predict that dyads who have a low degree of shared context (i.e., limited context) will be more inclined to align on forms, as they cannot make use of shared iconic form-meaning mappings for all objects. In addition, we think that low degrees of shared context will lead to a preference for iconic strategies that are visually shared (drawing and representing), as opposed to acting strategies which require training on shared 'cultural' information.

Method

Participants

Three experimental conditions were run. Participants were recruited on Prolific Academic. Participants reported to be fluent speakers of English and had not taken part in a previous study related to the present one (a related pilot study). All participants were paid 9 pounds for their participation in the study. The shared context condition was run separately from the baseline condition and limited context condition (see condition information below). In both the baseline condition and the limited context condition, 29 dyads out of 40 completed the experiment. In the shared context condition, 32 dyads out of 50 completed the experiment. The dyads who did not complete the study (usually because one member of the dyad did not complete the interaction phase within 60 seconds and was excluded) were not considered in the analysis.

Materials and procedure

The experiment was coded in Python using oTree, an open-source software created for interactive experiments (Chen, Schonger, & Wickens, 2016). This experiment received approval from the University of Amsterdam's humanities ethics committee. Prior to taking part in the study, participants received information about the experiment and gave their consent to partake in the study.

Stimuli in this experiment consist of 12 unfamiliar objects from the NOUN database (Horst & Hout, 2016) and short videos showing gestural descriptions of these objects. The gestural descriptions chosen correspond to iconic strategies outlined in Ortega and Özyürek (2020). Each object is paired with four gesture videos, which loop simultaneously and show four strategies for depicting the object: representing[2], drawing and two acting strategies. An example of one array can be seen in Figure 2. We chose two acting descriptions that we thought could plausibly be used to describe each object, and a pilot study confirmed that participants easily pair the gestural video description with the object description.

The experiment consists of three phases. The manipulation between the three conditions occurs only in phase 1. In phase 1 (training), participants are trained on pairings between pictures of objects and a written description of what the object could be used for (one or more acting descriptions). Figure 2 shows an example of a phase 1 trial for a dyad from the baseline condition and one member of a dyad from the limited context condition (A) and for a dyad in the shared context condition (B). In each condition, the training differs as follows:

- **Baseline condition**: participants are trained on the same acting description as their partner. One of the two acting descriptions is randomly chosen and is assigned to both participants for that object.
- **Limited context condition**: participants are trained on a different acting description from their partner.
- **Shared context condition**: participants are trained on both acting descriptions. In the shared context condition, the order of the acting descriptions is randomized per participant for each object.

After participants are shown the object-description pairings, they are instructed to select a gesture video from an array of four videos. Each participant is shown the 12 objects in a randomized order per dyad. The position of the four gestural description videos (top left, bottom left, top right or bottom right) is randomly assigned per participant.

In phase 2 (interaction), participants in each dyad take turns as director and matcher in an object selection task with the goal of successfully communicating about the objects they were previously trained on. As director, the participant sees an object and is instructed to select a gesture video from an array of four videos to be sent to the matcher. As matcher, the participant receives a gesture video from the director and is instructed to select which object from an array of four objects they think the director is referring to. Then, participants receive full feedback. In phase 2 there are 48 trials and the participants alternate as director and matcher each trial. Each participant is the director for every object twice and the matcher for every object twice. The director's gesture video array and the matcher's object array are randomized for each trial.

Finally, in phase 3 (recall), participants are asked for each object which gesture video was used for successful interaction with their partner. In all other aspects, the set-up is identical to that of phase 1.

Analysis

The data and analysis files can be found here: `https://figshare.com/s/d620f5ef9f181629841a`.
We focus our analysis on the difference between the limited context condition and the shared context condition. The baseline condition largely serves to confirm that participants choose the same gesture video when they are trained on the same acting description.

First, in our analysis, we assess if there is more lexical variation in the shared context condition over time (comparing the training and recall phases) compared to the limited context condition. We run a linear mixed effect model to investigate the effect of lexical distances in phase 1 and phase 3 as well as the effect of condition[3]. Following this, we conduct a series of exploratory analyses: we investigate if the reduction in variation can be explained by differences in error across conditions. Additionally, we look at subsets of the shared context condition which can be more accurately compared to

[2] In contrast to the definition of representing by Ortega and Özyürek (2020) which allows for movement in the hands and arms, our representing gestures are static.

[3] lmer(distance ˜ condition * phase + (1+condition+phase|object) + (1+phase|dyad)

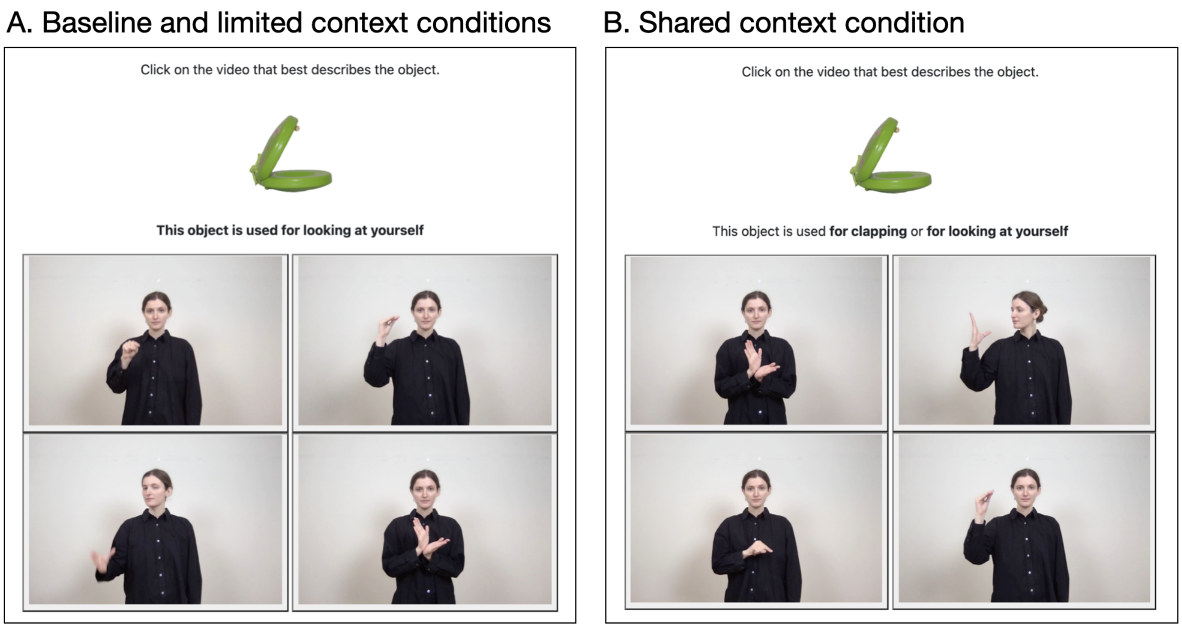

Figure 2: An example trial from phase 1 (improvisation). After having been shown the description associated with the object, participants are asked to select the pre-recorded gesture video that best describes the object. The four gesture videos correspond to the iconic strategies described in Ortega and Özyürek (2020): drawing in (top left A and bottom left B) , acting "for clapping" (top right A and bottom right B), acting "for looking at yourself" (bottom left A and top right B) and representing (bottom right A and top left B). In the baseline condition, both participants in the dyad would see the trial presented in A. In the limited context condition, one participant in the dyad would see the trial presented in A and their partner would see the other action description corresponding to this object "for clapping". In the shared context condition, both participants would see the trial shown in B. The gesture videos from Figure 2 can be found here: `https://figshare.com/s/80c71217a4ba12a97e98`

the limited context condition to see if they behave differently given their different training phases. Finally, we investigate if participants in the shared and limited context conditions have different preferences with regards to iconic strategies. We check if participants in the limited context condition increase their use of drawing and representing strategies in the recall phase. For each condition, we run a generalized linear mixed effects model to investigate the effect of phase (training vs. recall) on strategy (acting strategy vs. non-acting strategy)[4].

Results

We first look at differences in the amount of lexical variation across conditions, as shown in Figure 3. In the baseline condition, the mean lexical distance in the training phase (M=0.07; SD=0.25) is similar to that of the recall phase (M=0.08; SD=0.27). In the limited context condition, there is a mean lexical distance of 0.96 (SD=0.19) in the training phase which drops to 0.32 (SD=0.47) in the recall phase. In comparison, in the shared context condition, the mean lexical distance in the training phase is 0.52 (SD=0.5) which drops to 0.23 (SD=0.42) in the recall phase. We test if dyads in the shared context condition maintain more lexical variation than those in the limited context condition, and find that they do: a linear mixed effects model with an interaction between condition and phase explains significantly more of the variance than a model without the interaction ($X^2(1)=29.55$, $p<0.001$), showing that distance is reduced more strongly in the limited context condition.

Though the decrease in lexical distance is greater in the limited context condition compared to the shared context condition, the mean lexical distance in the recall phase is lower in the shared context condition than in the limited context condition, as can be seen in Figure 3. This is not what was expected from our hypothesis. In the limited context condition, participants are trained on different cultural descriptions per object and hence in the training phase they are nearly maximally different in their productions (M=0.96; SD=0.19). On the other hand, in the shared context condition, where participants were trained on both acting descriptions for each object, participants did not have a preference for one acting strategy over the other. Per dyad, this resulted in about half of trials where participants chose the same gesture video (lexical distance of 0) and about half of trials where participants chose a different gesture video (lexical distance of 1). This is reflected in the mean lexical distance starting close to 0.5 (M=0.52; SD=0.50). The current experimental design cannot

[4] glmer(use of non-acting strategies ~ phase + (1+phase|object) + (1+phase|dyad)

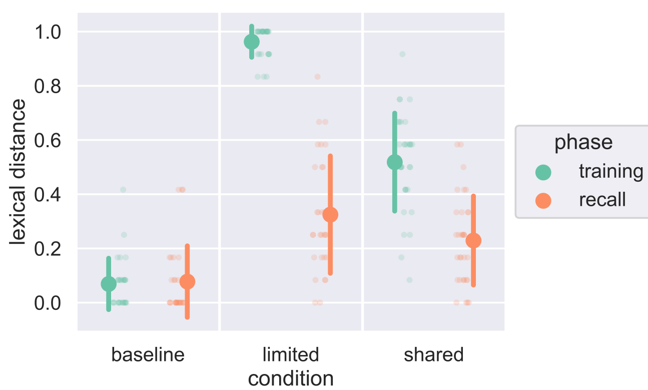

Figure 3: A dot plot showing the mean lexical distance of each condition by phase with error bars showing the standard deviation. Smaller dots show the mean of each dyad. A lexical distance of 0 indicates full alignment. Lexical distance is reduced significantly more over the course of the experiment in the limited context condition compared to the shared context condition.

rule out that in both the limited and shared context conditions participants *just* aim to reduce variation, and that in the limited context condition, participants simply have more initial variation to reduce (M=0.96; SD=0.19) compared to in the shared context condition (M=0.52; SD=0.50).

We now investigate why participants are reducing variation across conditions, starting with a focus on error rates in interaction. As a reminder, based on literature from experimental semiotics we would expect participants to align in order to overcome communicative errors. In the limited context condition, where participants were trained on different cultural information, we would expect error rates to be higher than in the other two conditions. This is what we found: error rates in the first half of the interaction phase of the shared context condition (M=0.06) are much more comparable to the baseline condition (M=0.04) than to the limited context condition (M=0.22). This aptly explains why lexical variation decreases in the limited context condition according to our theory (to overcome communicative errors), but why lexical variation is decreasing in the shared context condition (where there is much less error to overcome) requires more exploration.

We further assessed the decrease in lexical variation in the shared context condition by conducting two additional analyses. We aim to answer if the amount of variation observed in the recall phase of the shared context condition is a result of a) retaining variation because participants can make use of iconic affordances thanks to shared context (as hypothesized), or b) decreasing the amount of variation due to a general preference for variation reduction, independent of the role of shared context.

Our first exploration considers a subset of trials from the shared context condition where participants selected different gesture videos in the training phase (e.g., participant 1 selected ”for clapping” and participant 2 selected ”for looking at yourself” for the object shown in Figure 2). We compare these training phase trials (which have a lexical distance of 1) to those of the limited context condition (M=0.96; SD=0.19). If our hypothesis were to be valid, we would expect trials in the shared context condition that start out with a high degree of variation to retain that variation much more than in the limited context condition. Going against our hypothesis, what we find is more alignment in the subset of the shared context condition (M=0.30; SD=0.46) than in the limited context condition (M=0.32; SD=0.47).

Related to this, in our second exploration we entertain the possibility that dyads may act in a systematic way and should be analyzed as such (instead of being analyzed on a trial-by-trial basis). We make use of the fact that some dyads happen to have less lexical alignment initially than others. We subset the dyads in the training phase of the shared context condition who have a mean lexical distance above the average (M=0.52; SD=0.18). We compare these dyads to those in the limited context condition, whose lexical variation was nearly at ceiling in the training phase (M=0.96; SD=0.19). Going against our hypothesis that we should find more retention of variation in the shared context condition, we find that dyads in the shared context condition with an initially high degree of lexical variation have a lower degree of lexical variation in the recall phase (M=0.24; SD=0.15) compared to the dyads in the limited context condition (M=0.32; SD=0.47). In the discussion section, we will consider possibilities as to why our hypothesis was not borne out.

To assess if shared context influences the iconic strategies preferred by our participants, we compare the different iconic strategies in the shared vs. limited context conditions. In our experiment, visually shared iconic strategies are drawing and representing as their meaning can be retrieved visually from the form of the object. On the other hand, acting descriptions are culturally taught in the training phase, and provide the mapping between the gesture video and object. As such, we expected participants in the limited context condition to increase their use of visually shared iconic strategies (drawing and representing) as the use of acting is more likely to lead to communicative error. As participants in the shared context condition are trained on both acting descriptions, we predict that they will continue to use the acting strategy.

As expected, there is a strong initial preference for the acting strategy (because of training on these strategies): the acting strategy accounted for 94.68% of trials in the limited context condition and 92.83% of trials in the shared context condition. In the recall phase, as shown in Figure 4, this picture changes: participants in the limited context condition decrease their use of the acting strategy to 80%. The use of drawing increases from 2.15% in the training phase to 9.62% in the recall phase, and the use of representing increases from 3.16% in the training phase to 10.34% in the recall phase. As such, we find a significant increase in the use of non-acting strategies from the training phase to the recall phase

(β=1.31, SE=0.50, p<0.001), likely because the use of acting strategies led to communicative errors while non-acting strategies were iconic to both participants, relating to shared visual properties of the objects. Also in line with our prediction, in the shared context condition, there is no significant increase in the proportion of non-acting strategies (drawing and representing) from the training phase to the recall phase (β=-0.44, SE=0.46, p=0.34), presumably because participants communicate successfully using the acting strategies. The use of acting only decreases from 92.83% of gesture video choices in the training phase to 91.27% in the recall phase.

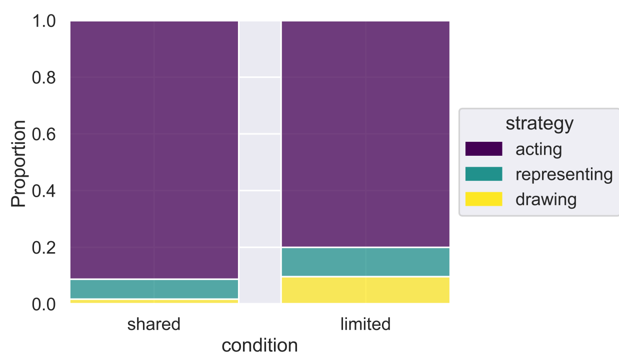

Figure 4: A stacked barplot showing the iconic strategies used in the recall phase in the shared and limited context conditions. In the limited context condition, the use of non-acting strategies (drawing and representing) significantly increases while the use of non-acting strategies in the shared context condition does not increase significantly.

Discussion

We have presented an experiment that allowed us to study how the amount of shared context and the iconicity of form-meaning mappings interact to influence lexical variation. We predicted that shared context would lead to the retention of the lexical variation initially present in a population and that limited context would encourage participants to align on a lexicon of visually shared iconic gestures. Comparing the differences between training and recall, we observed that participants in the limited context condition indeed aligned more strongly (resulting in a steeper downward slope of lexical distance). However, alignment levels in the recall phase were actually higher in the shared context condition, and this was unexpected. Additional analyses suggest that it was not errors in communication that led participants in the shared context condition to align, since the error rate in this condition was very low. This indicates a mismatch with the theory laid out in the modeling work (Mudd et al., 2022), which postulates that alignment is driven mainly by communicative error, and communicators do not align when they do not need to. This mismatch is a good illustration of why it is important to combine computational work with empirical testing: the simulations were useful in exploring the population level dynamics around alignment and iconicity, but our current experimental work provides evidence that the cognitive assumptions of the model need to be refined.

What may push participants in the shared context condition to align despite their ability to communicate successfully while retaining variation? It seems likely that participants in our experiment simply wish to align. First, it may be that participants have pragmatic assumptions that they also apply to experimental contexts (similarly to why participants probability match when they don't need to in Perfors, 2016). In addition, participants (speakers of English) are probably influenced by their real-life language experience which may influence their willingness to align with their partner. As such, linguistic norms, such as tolerance for variability, may vary in different contexts and it is possible that as speakers we have some meta-linguistic awareness which influences our inclination to align with our interlocutor. It would be interesting to further explore if there is a group of people (e.g., of a certain linguistic background, with certain personality traits) that would be happier maintaining variation in this setting.

Apart from assessing the effect of shared/limited context on lexical variation, our experiment allowed us to study the use of different iconic strategies. In line with our hypothesis, we found that participants in the limited context condition increased their use of non-acting strategies (drawing and representing). We posit this is because of the difference in epistemic properties between the acting and non-acting strategies: the non-acting strategies rely on purely visual iconic mappings (referring to the shape of the object) and require no cultural knowledge (of actions associated with the object). Participants in the limited context condition prefer to align on visually shared iconic gestures, likely because of communication errors when using the acting strategy. In other words, participants adapt their linguistic choices based on their linguistic niche (in line with Lupyan & Dale, 2016). This provides support for the view by Occhino et al. (2017) that "iconicity should be viewed as an accepted affordance used by all languages to varying degrees in varying linguistic contexts" (p.104). Here we have shown that communicators are aware of the amount of shared information and use this knowledge in choosing an iconic strategy.

Conclusion

In our experimental setup we were able to test the influence of shared context on lexical variation in iconic manual communication. Our results do not support findings in previous work that shared context allows for high levels of lexical variation; instead we found that despite already communicating successfully, participants with high levels of shared information still reduced lexical variation. Using pre-recorded gesture videos allowed us to directly compare different iconic strategies, and here we observed a crucial role for the amount of shared context: strategies that did not require specific 'cultural' knowledge were preferred when the amount of shared knowledge was low.

References

Aronoff, M., Meir, I., & Sandler, W. (2005). The paradox of sign language morphology. *Language*, *81*(2), 301.

Boyd, R., & Richerson, P. J. (1988). *Culture and the Evolutionary Process*. University of Chicago Press.

Brown, G., Anderson, A., Schillock, R., & Yule, G. (1984). *Teaching talk.* Cambridge: Cambridge University Press.

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97. doi: 10.1016/j.jbef.2015.12.001

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*(5).

Clark, H. H. (1996). *Using language*. Cambridge university press.

de Vos, C. (2011). Kata Kolok Color Terms and the Emergence of Lexical Signs in Rural Signing Communities. *The Senses and Society*, *6*(1), 68–76. doi: 10.2752/174589311X12893982233795

Emmorey, K. (2014, September). Iconicity as structure mapping. *Phil. Trans. R. Soc. B*, *369*(1651), 20130301. doi: 10.1098/rstb.2013.0301

Fay, N., Arbib, M., & Garrod, S. (2013). How to Bootstrap a Human Communication System. *Cognitive Science*, *37*(7), 1356–1367. doi: 10.1111/cogs.12048

Fay, N., Ellison, M., & Garrod, S. (2014). Iconicity: From sign to system in human communication and language. *Pragmatics & Cognition*, *22*(2), 244–263. (Publisher: John Benjamins) doi: 10.1075/pc.22.2.05fay

Fay, N., Garrod, S., Roberts, L., & Swoboda, N. (2010). The Interactive Evolution of Human Communication Systems. *Cognitive Science*, *34*(3), 351–386. doi: 10.1111/j.1551-6709.2009.01090.x

Fusaroli, R., & Tylén, K. (2016). Investigating Conversational Dynamics: Interactive Alignment, Interpersonal Synergy, and Collective Task Performance. *Cognitive Science*, *40*(1), 145–171. doi: 10.1111/cogs.12251

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*(1), 8–11. doi: 10.1016/j.tics.2003.10.016

Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behavior Research Methods*, *48*(4), 1393–1409. doi: 10.3758/s13428-015-0647-3

Hudson Kam, C. L., & Chang, A. (2009). Investigating the cause of language regularization in adults: Memory constraints or learning effects? *J Exp Psychol Learn Mem Cogn*, *35*(3), 815–821. doi: 10.1037/a0015097

Krauss, R. M., & Weinheimer, S. (1964). Changes in reference phrases as a function of frequency of usage in social interaction: a preliminary study. *Social behavior and Attitudes*, *1*(1-12), 113–114. doi: 10.3758/BF03342817

Lupyan, G., & Dale, R. (2016). Why Are There Different Languages? The Role of Adaptation in Linguistic Diversity. *Trends in Cognitive Sciences*, *20*(9), 649–660. doi: 10.1016/j.tics.2016.07.005

Lutzenberger, H., Mudd, K., Stamp, R., & Schembri, A. C. (2023). The social structure of signing communities and lexical variation: A cross-linguistic comparison of three unrelated sign languages. *Glossa: a journal of general linguistics*, *8*(1). doi: 10.16995/glossa.10229

Macuch Silva, V., Holler, J., Ozyurek, A., & Roberts, S. G. (2020). Multimodality and the origin of a novel communication system in face-to-face interaction. *Royal Society Open Science*, *7*(1), 182056. doi: 10.1098/rsos.182056

Meir, I., Israel, A., Sandler, W., Padden, C. A., & Aronoff, M. (2012). The influence of community on language structure: Evidence from two young sign languages. *Linguistic Variation*, *12*(2), 247–291. (Publisher: John Benjamins) doi: 10.1075/lv.12.2.04mei

Mesoudi, A. (2011). An experimental comparison of human social learning strategies: payoff-biased social learning is adaptive but underused. *Evolution and Human Behavior*, *32*(5), 334–342. doi: 10.1016/j.evolhumbehav.2010.12.001

Mudd, K., de Vos, C., & de Boer, B. (2022). Shared context facilitates lexical variation in sign language emergence. *Languages*, *7*(1), 31. doi: https://doi.org/10.3390/languages7010031

Müller, T. F., Winters, J., & Morin, O. (2019). The Influence of Shared Visual Context on the Successful Emergence of Conventions in a Referential Communication Task. *Cognitive Science*, *43*(9), e12783. doi: 10.1111/cogs.12783

Occhino, C., Anible, B., Wilkinson, E., & Morford, J. P. (2017). Iconicity is in the eye of the beholder: How language experience affects perceived iconicity. *Gesture*, *16*(1), 100–126. doi: 10.1075/gest.16.1.04occ

Ortega, G., & Özyürek, A. (2020). Systematic mappings between semantic categories and types of iconic representations in the manual modality: A normed database of silent gesture. *Behavior Research Methods*, *52*(1), 51–67. doi: 10.3758/s13428-019-01204-6

Perfors, A. (2016). Adult Regularization of Inconsistent Input Depends on Pragmatic Factors. *Language Learning and Development*, *12*(2), 138–155. doi: 10.1080/15475441.2015.1052449

Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a General Property of Language: Evidence from Spoken and Signed Languages. *Front. Psychology*, *1*. doi: 10.3389/fpsyg.2010.00227

Sperber, D., & Wilson, D. (1996). *Communication and cognition (2nd ed.)*. Oxford, UK; Cambridge, MA: Blackwell Publishers.

Steels, L. (1995). A Self-Organizing Spatial Vocabulary. *Artificial Life*, *2*(3), 319–332. doi: 10.1162/artl.1995.2.3.319

Tamariz, M., Ellison, T. M., Barr, D. J., & Fay, N. (2014). Cultural selection drives the evolution of human communi-

cation systems. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1788), 20140488.
Taub, S. F. (2001). *Language from the body: Iconicity and metaphor in american sign language*. Cambridge University Press.
Thompson, B., Raviv, L., & Kirby, S. (2020). Complexity can be maintained in small populations: a model of lexical variability in emerging sign languages. In *Proceedings of the 13th International Conference on the Evolution of Language* (p. 3). doi: doi:10.17617/2.3190925
Tkachman, O., & Hudson Kam, C. L. (2020). Measuring lexical and structural conventionalization in young sign languages. *Sign Language and Linguistics*, *23*(1-2), 208–232. doi: 10.1075/sll.00049.tka
Winters, J., Kirby, S., & Smith, K. (2018). Contextual predictability shapes signal autonomy. *Cognition*, *176*, 15–30.