Repeated Interactions Can Lead to More Iconic Signals

Hannah Little (hannah.little@mpi.nl), Marcus Perlman (marcus.perlman@mpi.nl)
Language and Cognition Department, Max Planck Institute for Psycholinguistics, 6500 AH Nijmegen, The Netherlands

Kerem Eryılmaz (kerem@ai.vub.ac.be)
Artificial Intelligence Laboratory, Vrije Universiteit Brussel, 1050 Brussels, Belgium

Abstract

Previous research has shown that repeated interactions can cause iconicity in signals to reduce. However, data from several recent studies has shown the opposite trend: an increase in iconicity as the result of repeated interactions. Here, we discuss whether signals may become less or more iconic as a result of the modality used to produce them. We review several recent experimental results before presenting new data from multi-modal signals, where visual input creates audio feedback. Our results show that the growth in iconicity present in the audio information may come at a cost to iconicity in the visual information. Our results have implications for how we think about and measure iconicity in artificial signalling experiments. Further, we discuss how iconicity in real world speech may stem from auditory, kinetic or visual information, but iconicity in these different modalities may conflict.

Keywords: Iconicity; Modality; Artificial Language Experiment; Communication; Conventionalisation

Introduction

Roughly 7000 languages are spoken around the world, and dozens more are signed. Over the course of human history, according to one rough estimate, hundreds of thousands of languages may have passed in and out of existence (Pagel, 2000). The number of words that have cycled through human languages, then, is enormous, perhaps in the order of billions. Imagine that we could trace these symbols back to their origins. How did people create the first words and signs?

One hypothesis is that the first words were created using *iconicity* (Fay, Ellison, & Garrod, 2014; Imai & Kita, 2014; Perlman, Dale, & Lupyan, 2015; Perniss, Thompson, & Vigliocco, 2010). Iconicity is a quality of a signal that, regardless of modality or medium, exhibits a degree of resemblance between its form and its meaning. For example, a person can communicate the idea of a 'rounded' shape by drawing a picture that resembles it, by molding their hands to reflect the shape, or by vocalising a 'round' word like "bouba". Iconicity can function to jump-start a new communication system because it enables a communicator to create new signals that are, to some extent, understandable to a partner without a shared system of conventional symbols.

The hypothesis that the first words were iconic improvisations is supported by evidence from natural signing systems. Traces of iconic creation are apparent in many of the signs of signed languages, and when signers lack a name for a referent, they tend to create an iconic sign for it (Klima & Bellugi, 1979). Further, when deaf children are raised without native signers and deaf peers, they create iconic gestures that ground the development of home sign systems that they use with hearing adults (Goldin-Meadow, 2003). Experimental studies where participants communicate using unfamiliar signalling systems also demonstrate extensive use of iconicity to ground novel signals, for example with drawing (Garrod, Fay, Lee, Oberlander, & MacLeod, 2007), slide whistles (Verhoef, Roberts, & Dingemanse, 2015), and non-linguistic vocalisations (Perlman et al., 2015).

In comparison to signed languages, the role of iconicity in the creation of spoken languages is obscure. It is widely assumed that spoken languages have markedly less iconicity than signed languages. Yet, it is unclear why this is the case. One widely argued reason is that the vocal-auditory modality affords little iconicity to represent a rich array of meanings (Armstrong & Wilcox, 2007). This argument is supported mainly by comparing impressions of the iconicity of gesture and sign with vocalisations and speech, and also by experimental studies finding that gestures were more effective than non-linguistic vocalisations at communicating different meanings (Fay, Arbib, & Garrod, 2013; Fay, Lister, Ellison, & Goldin-Meadow, 2014). A second possible reason that spoken languages have so little iconicity is their extremely ancient origins. Over so many generations, the original iconicity of spoken languages has mostly degraded. This alternative assumes a process of conventionalisation in which the high level of iconicity characteristic of novel signals decays uni-directionally over time until it eventually disappears.

Iconicity and conventionalisation

Is it actually the case that the iconicity of novel signals necessarily decays over time as the signal becomes conventionalised? In signed languages, the iconicity of signs does appear to fade over time as the forms become more regularised and systematic (Frishberg, 1975). Although mature signed languages are still iconic to a large extent, they are nevertheless much younger than spoken languages, and we do not know what might happen to their iconicity with further development. Graphic systems may provide a clearer case of how iconicity diminishes over time. For example, early records of written Sumerian, early Egyptian and ancient Chinese show that they originated from more detailed, iconic depictions that have became conventionalised into an increasingly abstract code (Gelb, 1952; Sampson, 1985; Vaccari & Vaccari, 1964). A smaller-scale, but comparable, process for graphic systems has been demonstrated in the laboratory where drawings lose their iconicity and become more symbolic and arbitrary over repeated interactions (Caldwell & Smith, 2012; Garrod et al., 2007; Theisen, Oberlander, & Kirby, 2010).

However, recent experimental studies have found that sig-

nals may sometimes gain iconicity over repeated interactions, even as they otherwise show evidence of conventionalisation. In Perlman et al. (2015), pairs of participants took turns over ten rounds creating non-linguistic vocalisations for different meanings (e.g. big, rough, up). Accuracy within the game increased to ceiling, and vocalisations showed signs of conventionalisation, becoming shorter in duration and more stable in form. To measure how iconicity changed over this process, they tested the ability of naïve listeners to guess the meaning of vocalisations from rounds 1, 5, and 10. Vocalisations from round 1 were guessed with the lowest accuracy, suggesting they were the least iconic, but in later rounds, vocalisations were guessed with higher accuracy. Verhoef, Kirby, and de Boer (2014) also found that signals increased in iconicity over repeated interactions and iterations. Participants used digital slide whistles to communicate different left or right facing animals. The results showed that participants only encoded the direction of these animals after 2 or 3 generations in an iterated chain.

On the surface, these findings may seem at odds with the idea that the function of iconicity is to bootstrap the formation of a conventional signal. How can signals become initially more iconic and then maintain their iconicity over time, even as they became more conventionalised? One explanation for this result is that the creation of iconic signals in vocalisation is more challenging than in modalities like drawing or gesture. Thus, partners may initially need to explore the signal space and negotiate their shared intuition for a meaningful vocalisation. Over interactions, as signals become streamlined, the strongly iconic features that are found to be effective in distinguishing its meaning tend to be enhanced, while more idiosyncratic features are shed.

Experiments

Stimuli

Stimuli for the experiments presented in this paper come from a previous experiment (Little, Eryılmaz, & de Boer, in press). In this experiment, participants produced signals for meanings varying in shape, colour, and texture, which were designed to have no shared features (explained in Little et al., in press). Figure 1 shows the 15 meanings used in the experiment. Theremin-like signals were created using a "Leap Motion" controller: an infrared sensor that detects hand position (see Eryılmaz & Little, 2016 for details of the paradigm). Participant's hand position determined the pitch of audio signals. Left to right hand-positions created low to high pitches respectively with a non-linear, exponential relationship between hand-position and pitch. Participants were given this audio feedback in real-time as they produced the signals and participants could not see each other as they produced signals. These signals were used because they share some qualities with speech: they are auditory, continuous and restrict the use of iconicity. At the same time, they are non-linguistic and so minimise possible interference from pre-existing linguistic knowledge and conventions.

Figure 1: The meanings used in the experiment.

The stimuli were created in two experimental conditions: an 'individual' condition, where one person produced signals and received their own signals in batches of 5, and a 'communication' condition where two participants took it in turns to produce and receive signals. When receiving signals, participants were asked to identify their referent from an array of 4 meanings. Feedback was given on the correct answer immediately after each response in both conditions.

The meaning space expanded throughout the experiment: by 5 meanings at a time in every block in the individual condition, and by 2 meanings at a time in the communication condition. In the communication condition, the meaning space only expanded once the participants had agreed on signals for existing meanings (by communicating them correctly twice).

For the experiments in this paper, signals from "early" in the individual condition were taken from the first phase (5 signals) and "late" signals were taken from the last phase (15 signals). In the communication condition, no pair managed to finish the experiment before time ran out, and so all of the data from the "last phase" in the current paper is referring to the last phase participants got to in their particular experiment. "Early" signals from the communication condition were for the first 2 meanings seen.

Experiment 1: Audio playback experiment

We conducted a playback experiment to examine how the iconicity of signals changed over repeated interactions in the experiment above. Naïve listeners, without knowledge of a signal's development, guessed the meanings of the signals produced in the individual and communication conditions at both the beginning and end of the game. We took listeners' ability to match the signal with its intended referents as a measure of iconicity. This method for measuring iconicity has been used previously in a number of studies (e.g. Garrod et al., 2007; Perlman et al., 2015). The experiment tested two hypotheses, though it should be noted that both hypotheses could work in tandem, or represent different stages of emergence of a communication system.

Hypothesis 1 The first hypothesis is that in the communication condition, repeated interaction between two participants

will lead to initial signals that are high in iconicity, but then become less iconic over interactions. This would follow the results of experiments such as Garrod et al. (2007), that used drawings. Their results also suggest that we should not see a loss of iconicity in the individual condition as conventionalisation requires interaction between communicators.

Hypothesis 2 The second hypothesis is that iconicity will go up in the communication condition, in line with the findings of Perlman et al. (2015). If iconicity is not present from the beginning, or is very idiosyncratic, then interaction may act as a way for signals to adapt to be more transparently iconic. However, without an interlocutor, one would not expect there to be a pressure for transparency in any iconicity present, meaning signals in the individual condition should also not increase in iconicity under this hypothesis.

Method

Procedure 391 participants were recruited on social media. Each participant was sent to a webpage which redirected randomly to one of several signal sets on its own webpage. A signal set was typically 15 signals. Signals were mp3 files which were playable by the participants by clicking on them. Under each mp3 file was a set of 4 images of possible meanings including the correct referent and 3 others chosen at random. Participants were asked to click on the meaning of the four that "you think the sound refers to". They could change their mind as many times as they liked, and their responses were only recorded after they pressed "submit" at the bottom of the page.

Results and Discussion

The following results are all produced using a linear mixed effects analysis, from accuracy data that had been binnned by meaning. We included time phase (early or late) and condition (individual or communication) of production as fixed effects. The intended image was controlled for as a random effect with by-meaning random slopes for the effect of time phase and condition. Likelihood ratio tests were used to compare the model against a null model that did not include the variable of interest. The condition in which the signals were produced – individual or communication – did not appear to affect the iconicity of the signals ($\chi^2(1) = 0.1,\ p = 0.74$). Listeners correctly matched signals with their referents with nearly the same level of accuracy in both conditions (around 35%). The time phase in which the signal was produced also did not significantly affect guessing accuracy ($\chi^2(1) = 2.3$, $p = 0.13$). However, there was a significant interaction between condition and time produced ($\chi^2(1) = 5.9$, $p = 0.015$). In the graph (Figure 2), we show that naïve listeners were much better at matching signals that were produced later in the communication condition. In the individual condition, the signals went down slightly in their iconicity, though this difference was not significant.

The results from the audio playback experiment suggest that the iconicity of signals created at the start of the original

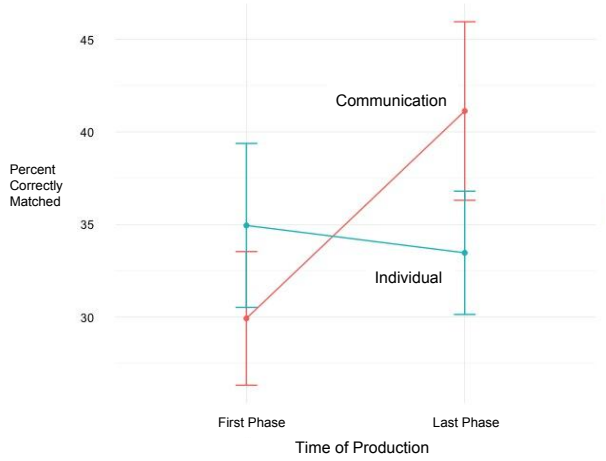

Figure 2: The percentage of signals correctly matched with their meanings by naïve listeners. The percentage for behaviour at chance levels is 25%.

communication game is nearly the same in both the individual and communication conditions. Naïve listeners were able to guess their meanings with nearly equal accuracy. This is not indicative of participants not attempting to be iconic (indeed, accuracy was above chance), but it may be that their attempts to be iconic start as being relatively idiosyncratic. In further support of this account, in the individual condition, the iconicity of signals did not change from the early to the late phase of the communication game. It may be that with an individual participant, there is no selection pressure to enhance the strongly iconic features of signals or to discard more idiosyncratic ones.

In contrast to the individual condition, we found that in the communication condition iconicity increased significantly from the early to the late phase, confirming hypothesis 2. In media that afford less iconicity, the presence of another person might cause ongoing pressure to enhance the iconicity of signals, making them increasingly transparent to naïve listeners. However, because of the multi-modal nature of the signals in the initial experiment (i.e. gesture generating audio signals), it is also a possibility that participants were just becoming accustomed to being iconic using the audio feedback (the only thing transmitted between participants), rather than using iconicity in the visual modality.

In the original study, Little et al. (in press) found that signals in the individual condition became longer and more complex over the course of the experiment. However, there was no evidence of signals changing in complexity in the communication condition. Within the experiment, participants were better at correctly matching signals with their correct referent in the individual condition (85.6% correct) than in the communication condition (74.4% correct). In the individual condition, participants improved at recognising the signals correctly throughout the experiment, but they got worse in the communication condition.

Experiment 2: Visual playback experiment

In this experiment, we look for evidence that participants were distinctly adapting the iconicity of their signals to be optimal for the communication modality. Many people's first instinct is to draw the shapes in the air, but this did not necessarily translate to an optimally iconic signal with respect to the auditory feedback that was generated. This auditory representation was the only information transmitted between participants as they could not see each other. Therefore, participants might have adapted their signals to be more iconic by sound, while at the same time discarding distracting features that turned out to be less iconic.

To examine more specifically how participants in the communication condition adapted their signals over the course of the experiment, we ran a second playback experiment where participants matched visual representations of the signals, instead of auditory ones. If signallers adapted their signals to enhance iconicity for the communication medium, but shed features that are less iconic, then naive guessing accuracy with the visual signals should not increase as it did for the auditory signals. Alternatively – as the visual signals include exactly that same information as the auditory signals, just mapped onto a spatial dimension – visual iconicity might increase along with auditory iconicity.

Method

Stimuli The stimuli were the same signals used in the audio playback experiment for the communication condition but transformed into visual representations. Signals were small (200x200px) videos of the hand trajectory used to produce the audio signals. A black square moves left and right in real time with how participants' hands moved to produce the signals. These videos were produced using only information from the x-axis of the hand trajectory. We only used information from the x-axis because only the x-axis affected the pitch of signals. This gave the naïve participants in the visual playback experiment the same amount of information as in the audio experiment, making them more directly comparable.

Procedure 97 participants were recruited on social media. Again, each participant was linked to a webpage that redirected them to a webpage with one of several possible signal sets. The procedure was the same as in Experiment 1, except that the stimuli were presented as videos instead of as audio files. Participants were asked to watch 15 videos each and choose the meaning that "you think the video refers to" for each signal.

Results and Discussion

We compared the results of the visual playback with the results from audio playback in Experiment 1. Again, these results are produced using a linear mixed effects analysis using data binned by meaning. For this experiment, time produced (early or late) and modality (audio or visual) were the fixed effects in the model. Meaning was controlled for as a random effect and the model had by-meaning random slopes for the effect of time phase and modality. Likelihood ratio tests were used to compare the model against a null model that did not include the variable of interest.

Guessing accuracy in both modalities is shown in figure 3. The modality of the signals – visual or auditory – did not affect the overall accuracy of selecting the correct image ($\chi^2(1) = 1.17$, $p = 0.28$). The time phase in which the signal was produced also did not significantly affect guessing accuracy across modalities ($\chi^2(1) = 1.4$, $p = 0.24$). However, there was a significant interaction between modality and time phase ($\chi^2(1) = 5.9$, $p = 0.015$). In the early phase, guessing accuracy was statistically equivalent in both modalities. However, in the later phase, while accuracy increased in the auditory condition, in the visual condition it dipped slightly (but not significantly).

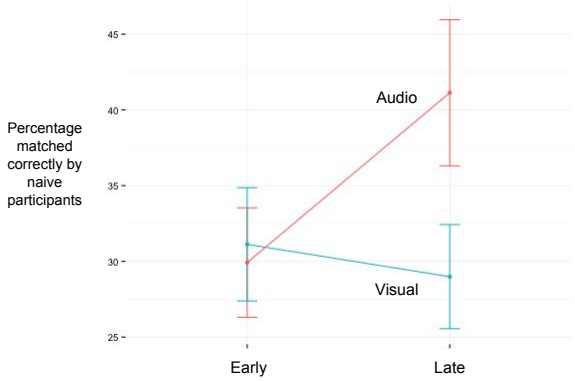

Figure 3: The percentage of both visual and audio signals from the communication game matched with their meanings by naïve participants. The percentage for behaviour at chance levels is 25%.

The results from the visual playback experiment demonstrate that signals produced at the beginning of the communication game exhibited a comparable level of iconicity in the auditory representations and the visual representations of the signals. However, as the iconicity appears to have increased in the auditory signals over the phases of the game, the iconicity of the visual transformations did not. This was the case even though the visual signals included the same information as the auditory signals. These results suggest that signallers in the communication game adapted their signals to be more iconic in ways that were particularly suited to the auditory communication channel. Features that may have been more iconic in a visual medium were not enhanced.

General Discussion

In the first playback experiment, we found that naïve listeners were more accurate at guessing the meanings of signals produced in the later phases of the experiment, but only in the communication condition. The pressure to became more iconic was only present when signals were being negotiated in interaction between individuals. One possible confound

here was that the meaning space expanding more quickly in the individual condition. Given that the signal space has only a limited amount of information to iconically encode meanings without ambiguity, it may be the the expansion meant a loss of iconic information across the whole meaning space. However, this consideration does not account for the fact that the meaning space also expanded in the communication condition where iconicity rose.

In a second playback experiment, we found that iconicity appeared to be enhanced particularly for the auditory communication medium, for which participants may have had weaker intuitions for iconicity compared to a visual medium. Together our findings demonstrate how, under certain conditions, the iconicity of signals can increase over repeated interactions, perhaps especially in a modality that affords less potential for iconicity. This may happen as partners initially explore the signal space and negotiate their shared intuition for a meaningful signal. Over interactions, as signals becomes streamlined with conventionalisation, the strongly iconic features may be agreed upon and enhanced, while more idiosyncratic features are back-grounded. This may only be an initial step in grounding a communication system, in running the experiment for longer, the signals may very well tend towards losing their iconicity again.

The results of this study have implications for semiotics experiments using artificial communication modalities and how iconicity is understood and measured in these studies. Many studies now use continuous auditory feedback as a result of some kinetic input, such as slide whistles (Verhoef, Kirby, & Boer, 2015), digital slide whistles (Verhoef, Roberts, & Dingemanse, 2015) and the Leap Motion paradigm (Eryılmaz & Little, 2016). Iconicity has been measured in signals generated from all of these paradigms, but not always in the same way. In Verhoef, Roberts, and Dingemanse (2015), the iconicity is measured by correlating the direction of stimuli (left or right facing animals) with the direction of pitch in a signal. Little, Eryılmaz, and de Boer (2015) measures iconicity by comparing the similarities between meanings with the similarities between signals, using information from the hand positions, rather than transformed values representing the auditory feedback. Verhoef, Kirby, and Boer (2015) asked naïve participants to rate how well signals "fit" the meanings they were paired with using auditory information alone. Importantly, none of these studies incorporate information from both the auditory and visual aspects of the signal in their measures for iconicity. Of course, there is a perfect correlation between movement and auditory feedback in all of these paradigms, but the results we present here suggest that iconicity may be perceived in very different ways depending on either the visual or auditory information. Some experiments using artificial continuous signal spaces do not have auditory feedback and are purely visual in nature (Galantucci, 2005; Verhoef, Walker, & Marghetis, 2016). These visual signals are treated as a proxy for a human communication systems in the same way that the paradigms with auditory feedback above are. However, it may be important to examine whether the results from such paradigms may, in some cases, be modality-specific.

Of course, caution is required in considering how our findings might generalise to languages and other natural communication systems. There are several reasons for reservation: the linguistic knowledge of our participants, the constrained signal and semantic space, the limited nature of the interaction, and the short time-scale of the experiment. Nevertheless, one interesting point of comparison may be the multimodality of our signals. In real-world communication, multimodality comes not only in the combination of speech and gesture, but also in the auditory and visual information that is conveyed by speech alone (Massaro, 1998; McGurk & MacDonald, 1976). This multi-modal nature of speech may impact how iconicity is encoded in speech. For instance, one common example of iconicity in spoken language is the /i/ phone for diminutive, as in words like teeny, itty-bitty (Ohala, 1994). This association has been found reliably across languages (Blasi, Wichmann, Hammarström, Stadler, & Christiansen, 2016). But what features of the /i/ make it iconic? Is it that the high pitch of the second and third formants corresponds with the high-pitched vocalisations of small animals? Is it the kinesthetic feel of articulating the sound, which is produced by contracting the oral cavity? Or might it be the visual features of the vowel, such as the speaker's retracted lips which resemble a submissive facial expression? These are difficult questions to answer, but future experiments might examine multi-modal signals and how iconicity is differentially informative across different modalities.

Further Work

The main reason for running the playback experiment with visual signals was the observation that participants were inclined to draw in the air as a starting point for novel signals. However, it is possible that this form of iconicity would only be evident from information from both the x- and y- axes. Though the y-axis did not affect the auditory feedback in any way, there was nothing to stop participants moving their hand vertically in the experiment. As a next step, we plan to create videos of participant's movements on both the x- and y- axes to see if such representations would exhibit a higher level of iconicity as a starting point, which might then decay.

Conclusion

In conclusion, we would like to challenge the oft-cited notion that languages consistently lose their iconicity over time. The work presented here and elsewhere (Perlman et al., 2015; Verhoef, Roberts, & Dingemanse, 2015) demonstrates the dynamic nature of iconicity in the evolution of symbol systems, which may adapt to the communication modality and the context in which it is used. Thus the multitude of morphemes cycling through languages may not always be drifting towards arbitrariness. In some cases, words and signs may become *more* iconic with time. The lexicons of natural languages, whether spoken or signed, exist in a bal-

ance between iconicity and arbitrariness (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Perniss et al., 2010).

Acknowledgements

Thanks to Gareth Roberts for feedback on the first playback experiment that lead to the implementation of the second. The original experimental work was financed by the ERC starting grant project ABACUS, grant number 283435.

References

Armstrong, D. F., & Wilcox, S. (2007). *The gestural origin of language*. Oxford University Press.

Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F., & Christiansen, M. H. (2016). Sound–meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences*, 201605782.

Caldwell, C. A., & Smith, K. (2012). Cultural evolution and perpetuation of arbitrary communicative conventions in experimental microsocieties. *PloS One*, *7*(8), e43807.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, *19*(10), 603–615.

Eryılmaz, K., & Little, H. (2016). Using leap motion to investigate the emergence of structure in speech and language. *Behavioral Research Methods*, doi:10.3758/s13428-016-0818-x.

Fay, N., Arbib, M., & Garrod, S. (2013). How to bootstrap a human communication system. *Cognitive science*, *37*(7), 1356–1367.

Fay, N., Ellison, M., & Garrod, S. (2014). Iconicity: From sign to system in human communication and language. *Pragmatics & Cognition*, *22*(2), 244–263.

Fay, N., Lister, C. J., Ellison, T. M., & Goldin-Meadow, S. (2014). Creating a communication system from scratch: gesture beats vocalization hands down. *Frontiers in Psychology*, *5*, 354.

Frishberg, N. (1975). Arbitrariness and iconicity: historical change in american sign language. *Language*, 696–719.

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science*, *29*(5), 737-767.

Garrod, S., Fay, N., Lee, J., Oberlander, J., & MacLeod, T. (2007). Foundations of representation: where might graphical symbol systems come from? *Cognitive Science*, *31*(6), 961–987.

Gelb, I. J. (1952). *A study of writing: The foundations of grammatology*. The University of Chicago Press.

Goldin-Meadow, S. (2003). *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language*. Psychology Press.

Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *369*(1651), 20130298.

Klima, E., & Bellugi, U. (1979). *The signs of language.* Cambridge: Harvard University Press.

Little, H., Eryılmaz, K., & de Boer, B. (2015). Linguistic modality affects the creation of structure and iconicity in signals. In D. C. Noelle et al. (Eds.), *The 37th annual meeting of the cognitive science society (cogsci 2015)* (p. 1392-1398). Austin, TX: Cognitive Science Society.

Little, H., Eryılmaz, K., & de Boer, B. (in press). Conventionalisation and discrimination as competing pressures on continuous speech-like signals. *Interaction Studies*.

Massaro, D. W. (1998). *Perceiving talking faces: From speech perception to a behavioral principle* (Vol. 1). Mit Press.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. , 746-748.

Ohala, J. J. (1994). The frequency codes underlies the sound symbolic use of voice pitch. In L. Hinton, J. Nichols, & J. J. Ohala (Eds.), *Sound symbolism* (pp. 325–347). Cambridge: Cambridge University Press.

Pagel, M. (2000). The history, rate and pattern of world linguistic evolution. *The evolutionary emergence of language: social function and the origins of linguistic form*, 391–416.

Perlman, M., Dale, R., & Lupyan, G. (2015). Iconicity can ground the creation of vocal symbols. *Royal Society open science*, *2*(8), 150152.

Perniss, P., Thompson, R., & Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. *Frontiers in psychology*, *1*, 227.

Sampson, G. (1985). *Writing systems: A linguistic introduction*. Stanford University Press.

Theisen, C. A., Oberlander, J., & Kirby, S. (2010). Systematicity and arbitrariness in novel communication systems. *Interaction Studies*, *11*(1), 14–32.

Vaccari, O., & Vaccari, E. E. (1964). *Pictorial chinese-japanese characters: a new and fascinating method to learn ideographs*. Vaccari's Language Institute.

Verhoef, T., Kirby, S., & Boer, B. (2015). Iconicity and the emergence of combinatorial structure in language. *Cognitive Science*.

Verhoef, T., Kirby, S., & de Boer, B. (2014). Emergence of combinatorial structure and economy through iterated learning with continuous acoustic signals. *Journal of Phonetics*, *43*, 57-68.

Verhoef, T., Roberts, S. G., & Dingemanse, M. (2015). Emergence of systematic iconicity: Transmission, interaction and analogy. In D. C. Noelle et al. (Eds.), *The 37th annual meeting of the cognitive science society (cogsci 2015)* (pp. 2481–2487). Austin, TX: Cognitive Science Society.

Verhoef, T., Walker, E., & Marghetis, T. (2016). Cognitive biases and social coordination in the emergence of temporal language. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *The 38th annual meeting of the cognitive science society (cogsci 2016)* (pp. 2615–2620). Austin, TX: Cognitive Science Society.