The benefits of one-sided iconicity

Petros Kaklamanis (pkakla@upenn.edu)
Department of Linguistics, University of Pennsylvania, USA

Robert Snider (rsnider8859@gmail.com)
Department of Theoretical and Applied Linguistics, University of Cambridge, UK

Gareth Roberts (gareth.roberts@ling.upenn.edu)
Department of Linguistics, University of Pennsylvania, USA

Abstract

Iconicity has recently been shown to be widespread in language and to play a particularly important role in bootstrapping new referring expressions or even getting whole new languages off the ground. The basis of this role has long been assumed to depend primarily on transparency for the receiver of the iconic signal, but might there also be producer-side advantages that play a significant role? We investigated this using an experimental referential communication game in which dyads communicated fruit and vegetables. We manipulated whether the sender could generate iconic signals and whether the receiver saw them. Results suggested that iconicity gave dyads a head-start, via stability in production, even if the receiver did not perceive the iconicity. However, this benefit declined over time, most likely due to memory constraints on the receiver.

Keywords: iconicity; communication; language; referential communication game; experimental semiotics; laboratory languages; audience design

Introduction

Over the last few decades, linguistic iconicity has increasingly become a focus of study. By *iconicity* we mean the phenomenon where communicative forms are motivated by what they refer to (e.g., onomatopoeia). While early work on language emphasized its arbitrariness (de Saussure, 1916/2013; Hockett, 1960), more recent work has argued that iconicity is in fact cross-linguistically pervasive (Dingemanse et al., 2015; Monaghan et al., 2014; Perniss et al., 2010).

Linguistic iconicity can be quite concrete, as with onomatopoeia and ideophones (Dingemanse, 2012), or quite abstract, as with reduplication, which often corresponds iconically to some kind of increase, such as plurality, iteration, or intensity (Downing & Stiebels, 2012).

Iconicity is particularly common in communicative situations which necessitate bootstrapping of new signals (Imai & Kita, 2014; Muysken, 2013). It features prominently in homesign systems and newly emerging sign languages (Cartmill et al., 2017; Coppola & Brentari, 2014; Sandler et al., 2011). It is also very widely observed in experiments involving novel non-linguistic communication systems (Roberts et al., 2015; Verhoef et al., 2016). An obvious reason for this is that iconicity aids comprehension by increasing semantic transparency for the audience (Auch et al., 2020). This comprehension benefit for novel forms aids the bootstrapping of new shared communication systems (Lockwood et al., 2016), a claim supported by evidence from experimental communication tasks and learning tasks (e.g., Perlman et al., 2015; Sato et al., 2020). But does iconicity also provide a distinct *production benefit* for the sender, such as by facilitating the generation of novel forms before they even reach the receiver? By providing a relatively fixed anchor point for signals, iconicity should also aid stability in production. While reference has been made in the literature to iconicity facilitating both the creation and comprehension of novel forms (e.g., Perlman et al., 2015), there is, to our knowledge, no experimental work directly separating the potential benefit of iconicity for the sender in generating novel forms from its direct benefit to the receiver in interpreting novel forms.

In this paper, we present an experimental investigation into this question using a *laboratory language game* paradigm (Galantucci et al., 2012; Nölle & Galantucci, 2022; Roberts, 2017) in which we manipulate whether iconicity is available to only the sender of a signal, both the sender and the receiver, or neither. This allows us to identify whether iconicity provides a communicative production benefit distinct from any comprehension benefit to the receiver.

Method

We designed a dyadic referential communication game in which a *Sender* generated signals to communicate referents (black-and-white images of fruits and vegetables) to a *Receiver*. In the *High-iconicity condition* the Sender could generate highly iconic signals (colors), which the Receiver would see. In the *Low-iconicity condition*, the Sender could instead generate a dot in 2D space, which the Receiver would again see. In the *One-sided iconicity condition*, the Sender could generate colors as in the High-iconicity condition, but—unknown to participants—the Receiver saw a dot instead. As such, the participants in this condition were misled about the communicative medium available to their partner, leading to an asymmetry in the form of the signals. We predicted that One-sided iconicity dyads would nonetheless communicate more successfully than Low-iconicity dyads despite (a) the asymmetry in signals, and (b) signals being no more iconic for the Receiver in either condition.

Participants

136 University of Pennsylvania students (10 left-handed), none of them colorblind, participated in dyads for course credit.

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

Materials

Physical equipment and software Participants sat in separate cubicles, from which they could not see each other, each with a computer (a mid-2014 Apple iMac with a 21.5" screen), running custom-designed software written in Python (Van Rossum & Drake, 1995) and Kivy (Virbel et al., 2011). The software was based on previous software used by Roberts and Clark (2020). Each participant wore a pair of noise-canceling headphones (specifically a Sony Wireless Noise Canceling Stereo Headset). Participants with the role of Sender used a 2012 Apple Magic Trackpad.

Experimental stimuli Referents consisted of a set of six black-and-white images of fruits and vegetables, selected—via a normalization study—as being relatively easy to communicate iconically using colors but not using a single dot.

Procedure

Participants played as Sender or Receiver in a dyadic cooperative communication game, never swapping roles. The game consisted of a series of rounds, in each of which the Sender saw a screen as in Figure 1. In the left half—the *referent panel*—six fruits and vegetables were arranged randomly, with the one in the center highlighted. This was the communication target that round, and it was selected randomly each round. The *message panel* and *sent-message panel* on the right of the screen appeared gray by default but would change in response to the Sender (See Experimental Conditions Section for details). Score and remaining round time were also displayed.

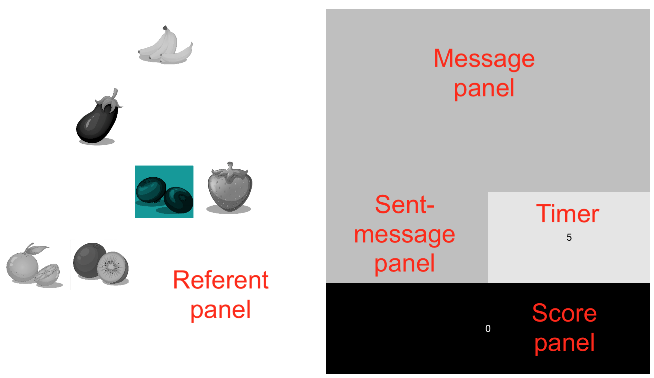

Figure 1: Screenshots of Sender's screen. Labels are for clarity and were not shown to participants.

The Receiver's screen was similar except that the referents were arranged in a different random order, with no referent in the center. The Receiver also had a cursor that could be moved around the referent panel using arrow keys.

The Sender's task was to convey the target by generating a signal with their trackpad (see Experimental Conditions Section). The Receiver's task, upon receiving this message, was to move their cursor to the correct referent and press the enter key to select it. Both players would then receive feedback. The Receiver was shown the target referent and the Sender was shown what the Receiver had selected, whether correct or not. If the selection was correct, the dyad scored a point. A round lasted max. 30 s but would end as soon as the Receiver had selected a referent. If the Receiver did not choose a referent within 30 s, the dyad scored no points.

Each game lasted 60 min in total unless the dyad successfully triggered the winning condition, which was to have successfully communicated *each* of the six referents on at least three of the previous four rounds in which it had been the target. As soon as this was achieved, the game would end. To incentivize good performance participants were told that the game would end early (with no reduction of reward) if they did "very well".

At the start of the experiment, participants played eight practice rounds that differed from the main rounds in three ways: First, they lasted 90 s rather than 30 s; second, the dyad's score did not carry over into the normal rounds; third, participants were reminded of their role at the start of each of the first two rounds. After the practice rounds ended, both participants were reminded of key points in the instructions.

Experimental Conditions

There were three between-subjects experimental conditions: *High-iconicity*, *One-sided iconicity*, and *Low-iconicity*. 22 dyads participated in the High- and Low-iconicity conditions each, while 24 dyads participated in the One-sided iconicity condition.[1] These conditions differed solely in the nature of the signals as they appeared to the Sender and Receiver. All other mechanics remained the same.

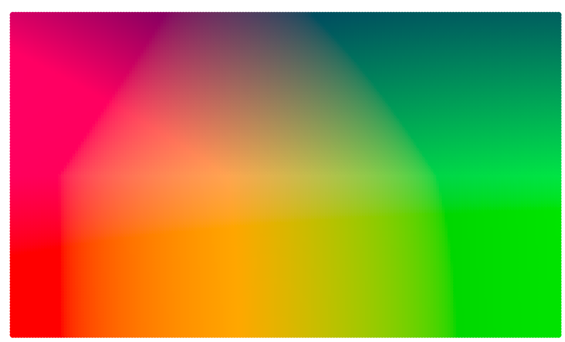

Figure 2: Underlying colorspace for high-iconicity and one-sided iconicity conditions

High-iconicity In the High-iconicity condition, the Sender's finger position on the trackpad would correspond dynamically to a color displayed in the message panel on the top right of the Sender's screen; this would update continuously based on finger position (the underlying colorspace is displayed in Figure 2). For instance, if the Sender positioned their finger in the bottom-left corner of the trackpad, the

[1]Owing to an oversight, two extra dyads were accidentally recruited for this condition.

message panel would turn red and would become more orange if they moved their finger rightwards. If the Sender lifted their finger off or touched the trackpad with more than one finger, the message panel would return to its gray default appearance. If they held their finger in place on the trackpad for at least 1 s, the currently displayed color would appear on both the Receiver's message panel and the Sender's sent-message panel. The Sender was told to move a finger around the trackpad and observe the screen, and to hold a finger in the same position for 1 s to send a message; otherwise they were not instructed how to use the signaling medium.

Low-iconicity The Low-iconicity condition worked the same except that the Sender's finger movements did not produce colors but instead corresponded to a circular black dot on the message panel. For instance, if the Sender had their finger in the bottom-left corner of the trackpad, the dot would appear in the bottom-left corner of the message panel, which could be sent to the Receiver as a signal.

One-sided iconicity In the One-sided iconicity condition, the Sender's experience was precisely as in the High-iconicity condition but, when they sent their signal, the Receiver would see a dot (as in the Low-iconicity condition) instead of the color the Sender saw and selected. Neither the Sender nor the Receiver was informed of this communicative mismatch.

Results

We begin by presenting our initial planned analysis, the results of which suggest compelling support for our hypothesis. However, having conducted this analysis we realized that—owing to a feature of our design—these results exaggerated the score for One-sided iconicity relative to High-iconicity. We thus follow that analysis with a revised analysis that corrects for this and suggests a more nuanced picture, in which one-sided iconicity provided stability of reference that gave a clear head-start to dyads in establishing signals, but which faded over time as dyads found it harder to establish signals for all six referents, likely due to memory constraints on the part of the receiver.

Initial success-index analysis

We calculated dyads' success following the method employed by Roberts and Clark (2020), which was designed to take into account the number of referents that participants established reliable signals for and how fast they did so. By "reliable signal," we mean that for at least three out of the four most recent occasions the referent in question was a target, the dyad communicated it successfully. It also took speed into account, as a dyad who establishes reliable signals for all referents within 100 rounds is doing better than one who takes 200 rounds. First, for each dyad, we identified how many signals they had established in each round of the game (not including practice rounds). Then we calculated a success index as $(\sum_1^{n_r} s)/6n_r$, where n_r is the number of rounds and the numerator is thus a cumulative count of s, the number of successfully established signals for referents in a given round (with six being the maximum possible given the number of referents).[2]

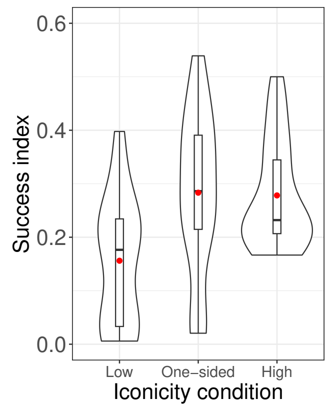

Figure 3: Violin plots of success index by condition, overlaid with boxplots. Practice rounds are not included. Red dot indicates mean.

Figure 3 shows success index by condition. Participants did best in the High-iconicity condition and worst in the Low-iconicity condition. The mean for One-sided iconicity was very close to the High-iconicity mean, and higher than in Low-iconicity, suggesting that iconicity does indeed provide a substantial communicative advantage, in line with our hypothesis. A linear model with success index as dependent variable and iconicity condition as predictor (with Low-iconicity as the reference level) supports this, with $\beta = 0.13, SE = 0.04, t = 3.43, p = 0.001$ for One-sided iconicity and $\beta = 0.12, SE = 0.04, t = 3.22, p < 0.001$ for the High-iconicity condition.

A revised success-index analysis

However, there is reason to think that this analysis downplays dyads' success in the High-iconicity condition and overplays their success in the One-sided iconicity condition. Figure 4 shows total game time by condition. As can be seen, dyads in the High-iconicity condition completed the game much faster than dyads in other conditions—no dyad took more than 19 minutes, and the mean was 10 minutes, a third of the mean time taken in the other two conditions! This is because—unlike in Roberts and Clark's (2020) experiment—the game ended early once dyads had established signals for all six referents. If dyads did not manage this, it continued for a full hour. In other words, dyads in the High-iconicity condition did much much better at the communication game than dyads in other conditions, and this is poorly represented in Figure 3.

Part of why the success index is underestimated for this condition is that it is based on data excluding the practice rounds. In earlier work employing this measure (in which the

[2] As noted by Roberts and Clark (2020), no player could actually score 1, as that would require them to have successfully communicated every referent several times before the start of the game. Correcting for this would have needlessly complicated an index intended as a relative measure.

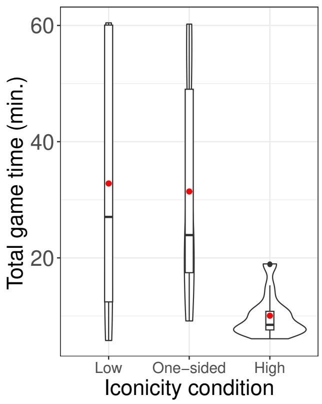

Figure 4: Violin plots of total game time by condition, overlaid with boxplots. Red dot indicates mean.

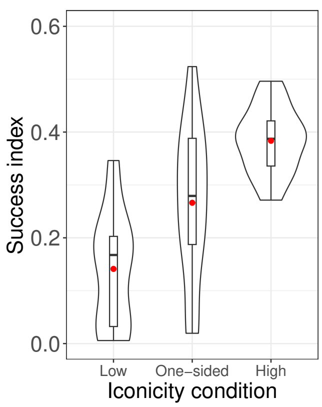

Figure 5: Violin plots of success index by condition, overlaid with boxplots. Practice rounds are included. Red dot indicates mean.

game never ended early) it made no difference to results if practice rounds were included or not. In our High-iconicity condition, however, dyads established signals so fast that the practice rounds took up a nontrivial proportion of their total rounds. Figure 5 shows success indices if practice rounds are included. As can be seen, the Low and One-sided Iconicity conditions are not much affected, but the success index for the High-iconicity condition is 37.9% higher than before.

The fact that—unlike in earlier work employing this measure—the game ended early if dyads did well also plays an important role. We can normalize for this by adding dummy rounds. This would simulate what would have happened if dyads had been made to continue playing for a full hour even if they established signals for all referents. We therefore calculated the mean length of the final four rounds of each game and—based on this—how many dummy rounds to add. Figure 6 shows the resulting success indices, making clear how well dyads in the High-iconicity condition did. However, while the success index for *this* condition remains significantly higher than for the Low-iconicity

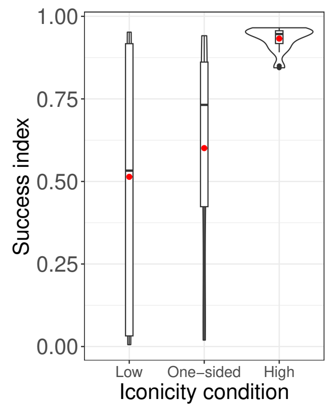

Figure 6: Violin plots of success index by condition, overlaid with boxplots. Practice rounds and dummy rounds are included. Red dot indicates mean.

condition, $\beta = 0.42, SE = 0.09, t = 4.59, p < 0.001$, the success index for the One-sided iconicity condition is no longer significantly higher than in the Low-iconicity condition: $\beta = 0.09, SE = 0.09, t = 0.97, p = 0.334$.

Round-by-round analysis

But this does not mean the difference suggested by Figures 3 and 5 is entirely illusory or meaningless. Dyads in the One-sided iconicity condition were in fact overall somewhat more accurate than in the Low-iconicity condition: $\beta = 0.1, SE = 0.05, t = 2.12, p = 0.038$. More importantly, however, they seem to have found it easier than dyads in the Low-iconicity condition to start establishing signals. We can investigate this by taking a closer round-by-round look at the basic data (including practice rounds). We ran a poisson linear mixed effects model on this data with number of reliable signals (as defined for the purpose of the success index) as the dependent variable, and round number (scaled) and condition (with Low-iconicity as the reference level) as independent variables, along with their interaction, and random intercepts for dyads. For One-sided iconicity, this revealed a main effect of condition, $\beta = 0.47, SE = 0.13, z = 3.63, p < 0.001$, and an interaction between round number and condition: $\beta = -0.27, SE = 0.03, z = 10.1, p < 0.001$.

The difference ultimately seems to come down to the ease of getting a signal set off the ground. Figure 7 shows, for each condition, the proportion of rounds dyads spent with different numbers of established signals. (Practice rounds, but not dummy rounds, are included.) As can be seen, most rounds ($> 75\%$) in the Low-iconicity condition were spent having established no signals at all. This was true for less than 48% of rounds in the other two conditions, and in this respect One-sided iconicity looks more like High-iconicity, the main difference being that dyads in the former spent more time (21% vs. 8%) with only one referent. Notably, however, dyads in One-sided iconicity spent more rounds (31.8%) with established signals for at least half the referents than Low-iconicity

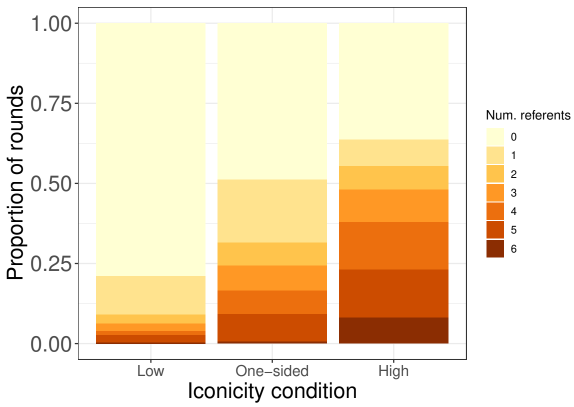

Figure 7: Proportion of rounds with each signal set size across all dyads, organized by condition (practice rounds included).

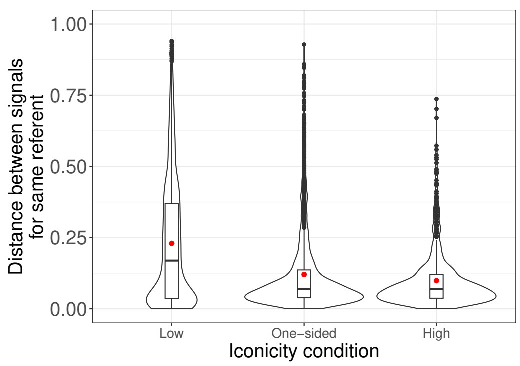

Figure 8: Violin plots of distances between successive signals for the same referent by condition, overlaid with boxplots. Practice rounds are included. Red dot indicates mean.

dyads spent with any established signals at all (21.7%).

In other words, One-sided iconicity dyads experienced a headstart. This seems to have been achieved via greater *signal stability*. We operationalized this by measuring, for each signal in each game, the distance between the Sender's finger position on the pad and their finger position for the previous signal for the same referent (Figure 8). Lower distances mean Senders were more consistently hitting the same spot. A linear mixed-effects model with distance as dependent variable, condition as independent variable (low-iconicity as reference level), and random intercepts for dyad and referent revealed an effect for both One-sided iconicity, $\beta = -0.07, SE = 0.02, t(65.85) = -3.1, p = 0.0029$ and High iconicity, $\beta = -0.09, SE = 0.02, t(71.82) = -3.8, p = 0.0003$. The same model with High iconicity as the reference level indicated no difference between that and One-sided iconicity: $\beta = -0.02, SE = 0.02, t(71.83) = -0.9, p = 0.37$

As suggested by the total game times, however (Figure 4), One-sided iconicity dyads did not find it so easy to *complete* the signal set. This likely reflects memory load on the Receiver's part: Memorizing one to three arbitrary signals was not too hard; memorizing six seems to have been rather more challenging.

Discussion

We conducted an experiment to investigate whether benefits of iconicity in bootstrapping a communication system are reducible to comprehension benefits for the receiver or if the dyad also gains from production benefits afforded to the sender. The experiment involved a Sender communicating by moving their finger on a trackpad to generate signals. In the High-iconicity condition signals consisted of colors, which both participants saw. In the Low-iconicity condition each signal consisted of a dot representing the Sender's finger position. In the One-sided iconicity condition the Sender produced colors, but the Receiver saw their finger position instead. Participants were unaware of this difference in experience.

High-iconicity dyads were very successful and fast at establishing a communication system; Low-iconicity dyads were the least successful. Dyads in the One-sided iconicity condition, however, were somewhat more accurate than Low-iconicity dyads, were much more stable in producing signals, and spent more rounds of the game having established signals for referents. In other words they experienced a head-start. However, this advantage declined over time, most likely due to pressures on the receivers' memory. Future work could investigate this possibility by varying the number of referents to be communicated, for instance, or having the Receiver complete a post-experiment recall test.

These results also have implications for alignment in communication. It is often assumed that successful communication relies on alignment between interlocutors at various levels (e.g., Garrod & Pickering, 2009; Poesio & Rieses, 2010). In our asymmetric One-sided iconicity condition, however, participants were misaligned in terms of the signals they saw and unaware of it. They nonetheless experienced an advantage over Low-iconicity dyads, who did not experience the same asymmetry. This is not to suggest that communicative alignment is not important but it does suggest that there is considerable flexibility in the kind of alignment needed to support successful communication—signal symmetry, in particular, may not be crucial—and that other factors, such as stability of reference, might be more important. It would be interesting in future work to disentangle the two factors and look at the role of asymmetric but similarly iconic signals.

The paper also has consequences for analysis. Our initial planned analysis is that reported at the beginning of the Results Section, in which we applied our success-index metric without including dummy rounds. While conducting this analysis and examining our data we realized that there was an important difference between our data and the data the metric had previously been used on. By comparing the results of our first analysis approach with subsequent analyses, we realized that the discrepancy indicated a more nuanced, interesting, and enlightening story than was initially apparent.

It is also important to note that the game involved a somewhat unusual and non-linguistic communication task. This is now a well-established and relatively widely used approach (Nölle & Galantucci, 2022) and the particular communication system was taken from an established paradigm (Roberts & Clark, 2020, 2023). Indeed, this approach was crucial to our investigation, as it allowed us to manipulate signal iconicity. However, it would be interesting in future work to investigate the role of unfamiliarity with the communication medium. One possibility is that the challenge of engaging with a novel communication medium might have increased participants' cognitive load and made the task of memorizing arbitrary signals (for Receivers in the One-sided iconicity condition and for both members of the dyad in the Low iconicity condition) particularly hard.

Such minor limitations aside, however, we consider that this study shines a nuanced light on the role of iconicity in bootstrapping signals and communication systems, and the basis of the communicative benefit it provides, while also contributing to research on the role of alignment in communication. It suggests that iconicity gives communicators a head-start in bootstrapping even if the audience doesn't know it's there.

Acknowlegements

We are grateful to several members of the Cultural Evolution of Language Lab for feedback and assistance, as well as to the Penn Undergraduate Research Mentoring program, and the National Science Foundation, which funded a project of which this is an offshoot (award number 1946882).

References

Auch, L., Gagné, C. L., & Spalding, T. L. (2020). Conceptualizing semantic transparency: A systematic analysis of semantic transparency measures in English compound words. *Methods in Psychology*, *3*, 100030.

Cartmill, E. A., Rissman, L., Novack, M. A., & Goldin-Meadow, S. (2017). The development of iconicity in children's co-speech gesture and homesign. *Language, Interaction and Acquisition*, *8*(1), 42–68.

Coppola, M., & Brentari, D. (2014). From iconic handshapes to grammatical contrasts: Longitudinal evidence from a child homesigner. *Frontiers in Psychology*, *5*, 830.

de Saussure, F. (1916/2013). Nature of the linguistic sign (R. Harris, Trans.). In, *Course in general linguistics* (pp. 75–81). Bloomsbury.

Dingemanse, M. (2012). Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass*, *6*(10), 654–672.

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.

Downing, L. J., & Stiebels, B. (2012). Iconicity. *The morphology and phonology of exponence*, 379–426.

Galantucci, B., Garrod, S., & Roberts, G. (2012). Experimental semiotics. *Language and Linguistics Compass*, *6*(8), 477–493.

Garrod, S., & Pickering, M. J. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science*, *1*(2), 292–304.

Hockett, C. F. (1960). The origin of speech. *Scientific American*, *203*(3), 88–97.

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

Lockwood, G., Hagoort, P., & Dingemanse, M. (2016). How iconicity helps people learn new words: Neural correlates and individual differences in sound-symbolic bootstrapping. *Collabra*, *2*(1), 7.

Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014). How arbitrary is language? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1651), 20130299.

Muysken, P. (2013). Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and cognition*, *16*(4), 709–730.

Nölle, J., & Galantucci, B. (2022). Experimental Semiotics: Past, present, and future. In A. M. García & A. Ibáñez (Eds.), *The routledge handbook of semiosis and the brain* (pp. 66–81). Routledge.

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. L., & Vigliocco, G. (2010). Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology*, *1*, 227.

Poesio, M., & Rieses, H. (2010). Completions, coordination, and alignment in dialogue. *Dialogue & Discourse*, *1*(1).

Roberts, G. (2017). The linguist's Drosophila: Experiments in language change. *Linguistics Vanguard*, *3*(1), 20160086.

Roberts, G., & Clark, R. (2020). Dispersion, communication, and alignment: An experimental study of the emergence of structure in combinatorial phonology. *Journal of Language Evolution*, *5*(2), 121–139.

Roberts, G., & Clark, R. (2023). The emergence of phonological dispersion through interaction: An exploratory secondary analysis of a communicative game. *Frontiers in Psychology*, *14*, 1130837.

Roberts, G., Lewandowski, J., & Galantucci, B. (2015). How communication changes when we cannot mime the world: Experimental evidence for the effect of iconicity on combinatoriality. *Cognition*, *141*, 52–66.

Sandler, W., Aronoff, M., Meir, I., & Padden, C. (2011). The gradual emergence of phonological form in a new language. *Natural Language & Linguistic Theory*, *29*, 503–543.

Sato, A., Schouwstra, M., Flaherty, M., & Kirby, S. (2020). Do all aspects of learning benefit from iconicity? Evidence from motion capture. *Language and Cognition*, *12*(1), 36–55.

Van Rossum, G., & Drake, F. L. (Eds.). (1995). *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam.

Verhoef, T., Kirby, S., & De Boer, B. (2016). Iconicity and the emergence of combinatorial structure in language. *Cognitive Science*, *40*(8), 1969–1994.

Virbel, M., Hansen, T., & Lobunets, O. (2011). Kivy–A framework for rapid creation of innovative user interfaces.