Cognition 141 (2015) 52–66

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

How communication changes when we cannot mime the world: Experimental evidence for the effect of iconicity on combinatoriality

Gareth Roberts [a,b], Jirka Lewandowski [a], Bruno Galantucci [a,c,*]

[a] Department of Psychology, Yeshiva University, New York, NY, USA
[b] Department of Linguistics, University of Pennsylvania, Philadelphia, PA, USA
[c] Haskins Laboratories, New Haven, CT, USA

ARTICLE INFO

Article history:
Received 10 March 2014
Revised 27 March 2015
Accepted 1 April 2015
Available online 30 April 2015

Keywords:
Experimental Semiotics
Signs
Language evolution
Bootstrapping communication
Sign language

ABSTRACT

Communication systems are exposed to two different pressures: a pressure for transmission efficiency, such that messages are simple to produce and perceive, and a pressure for referential efficiency, such that messages are easy to understand with their intended meaning. A solution to the first pressure is combinatoriality – the recombination of a few basic meaningless forms to express an infinite number of meanings. A solution to the second is iconicity – the use of forms that resemble what they refer to. These two solutions appear to be incompatible with each other, as iconic forms are ill-suited for use as meaningless combinatorial units. Furthermore, in the early stages of a communication system, when basic referential forms are in the process of being established, the pressure for referential efficiency is likely to be particularly strong, which may lead it to trump the pressure for transmission efficiency. This means that, where iconicity is available as a strategy, it is likely to impede the emergence of combinatoriality. Although this hypothesis seems consistent with some observations of natural language, it was unclear until recently how it could be soundly tested. This has changed thanks to the development of a line of research, known as Experimental Semiotics, in which participants construct novel communication systems in the laboratory using an unfamiliar medium. We conducted an Experimental Semiotic study in which we manipulated the opportunity for iconicity by varying the kind of referents to be communicated, while keeping the communication medium constant. We then measured the combinatoriality and transmission efficiency of the communication systems. We found that, where iconicity was available, it provided scaffolding for the construction of communication systems and was overwhelmingly adopted. Where it was not available, however, the resulting communication systems were more combinatorial and their forms more efficient to produce. This study enriches our understanding of the fundamental design principles of human communication and contributes tools to enrich it further.

1. Introduction

Emerging communication systems face two substantial challenges. One (which Hockett, 1960b, referred to as the problem of emission and detection) is to guarantee that a message is emitted accurately by its senders and detected

* Corresponding author at: Department of Psychology, Yeshiva University, New York, NY, USA.

E-mail address: bruno.galantucci@yu.edu (B. Galantucci).

http://dx.doi.org/10.1016/j.cognition.2015.04.001
0010-0277/

accurately by its receivers. We will refer to this as the challenge of *transmission efficiency*, as it is concerned with the transmission of messages through a noisy channel. The other challenge (which Hockett referred to as the problem of encoding and decoding) is to establish reference, such that once a message is detected, it is understood with the intended meaning. We will refer to this as the challenge of *referential efficiency*.

A solution to this latter challenge is to adopt communicative forms which are *iconic*, that is, which are not fully arbitrary, but are instead intuitively motivated by what they refer to (Burling, 1999; Donald, 1991). Iconic forms provide "scaffolding for the cognitive system to connect linguistic form and embodied experience" (Perniss, Thompson, & Vigliocco, 2010, p. 12), making it easier to establish new signs (Fay, Arbib, & Garrod, 2013; Fay, Lister, Ellison, & Goldin-Meadow, 2014).

A solution to the challenge of transmission efficiency is to adopt a design feature — *combinatoriality* — by which small sets of meaningless forms (such as phonemes or letters in English) are recombined to express an infinite number of meanings (Abler, 1989; Hockett, 1960a; Hockett, 1960b; Martinet, 1960; Nowak, Krakauer, & Dress, 1999; Studdert-Kennedy, 2000). Because these forms can be "chosen so as to be easily emitted and so as to be easily distinguished by the sensory receptors" (Hockett, 1960b, p. 421), combinatoriality greatly simplifies the task of transmitting signals through a noisy channel.

Two points should be borne in mind concerning the use of the term combinatoriality in this paper. First, it is important to clearly distinguish combinatoriality from *compositionality*, the recombination of meaningful forms (e.g., morphology and syntax in natural language). While compositionality implies a systematic mapping between meanings and signals (Krifka, 2001), combinatoriality implies no such systematic mapping, and the basic combinatorial forms (as in the case of phonemes) are meaningless. Although the two kinds of structure may emerge together, it is not clear that the factors involved in their emergence are likely to be the same. Indeed, as meaning–motivated structure, compositionality can be seen as an abstract form of iconicity, constituting a solution to the challenge of referential efficiency (cf. Tria, Galantucci, & Loreto, 2012).

The second point is that combinatoriality and iconicity are better treated as continuous variables than all-or-nothing features. In the former case, a communication system in which a small set of forms recur a great deal can be considered more combinatorial than a system in which a large set of forms recur very little. Morse code, for instance, exhibits extremely high combinatoriality, relying as it does on the recombination of only two basic meaningless forms. At the other extreme, Al-Sayyid Bedouin Sign Language (ABSL) exhibits almost no combinatoriality (Sandler, Aronoff, Meir, & Padden, 2011). With respect to iconicity, non-arbitrariness can range from complete transparency to highly opaque motivation (one of the two British Sign Language signs for "German", for instance, is recognizably iconic only if one knows that the Prussian military used to wear spiked helmets).

Given the difference between ABSL and Morse Code with respect to combinatoriality, it is notable that iconicity is extremely widespread in the former, but essentially absent in the latter. Indeed, based on these two examples, it might seem that iconicity and combinatoriality were entirely incompatible. Such a conclusion would be too strong, however. Most languages exhibit both combinatoriality and iconicity to some extent, and phonologically regular onomatopoeic words, such as English "cock-a-doodle-doo", illustrate that the two features can coexist in the same referring expression. Nevertheless, the mutually exclusive relationship between iconicity and combinatoriality exhibited in ABSL and Morse Code may be an indication that the two features are not fully independent, a possibility proposed by Sandler et al. (2011) to explain the absence of combinatoriality in ABSL. This lack of independence can be understood if one takes into account the requirements of an ideally combinatorial system and the requirements of an ideally iconic one. To best satisfy the pressure for transmission efficiency, combinatorial forms should be few in number (to maximize distinctiveness), simple to produce (to minimize production error, and because the smaller the set, the more frequently a given form is likely to be used), and lacking in independent meaning (otherwise recombination is limited, and larger numbers of forms are required). By contrast, the basic forms of an ideally iconic communication system are meaningful by definition, high in number (because they can be recombined only if the meaning fits), and relatively complex (to maximize distinctiveness[1]). The combinatorial iconic expressions found in languages like English constitute a compromise between these requirements (if it were not constrained by phonology, "cock-a-doodle-doo" might sound more like a real cock-crow). At the extreme ends of the continua, however, no compromise is possible. On the one hand, while iconicity in Morse code is certainly imaginable, the scope for it is extremely limited. On the other hand, if forms are iconic to the point of isomorphism with their referents, then combinatoriality is essentially precluded. That is, if communicative form is governed entirely by referent form, then whatever combinatorial structure a system might appear to have cannot be due to any organizational principle of the system itself, but must simply be a reflection of the structure of the meaning space captured by the system (and is thus at most compositional, not combinatorial).

That combinatoriality and iconicity might be in competition (or at least complementary distribution) was hinted at by Goldin-Meadow and McNeill (1999), who wrote:

> "the oral modality assumed the segmented and combinatorial code not because of its strengths but to compensate for its weaknesses. The oral modality is not well suited to conveying messages mimetically [i.e., iconically], even though that function is also important to human languages. This function is, however, very well served by the manual modality" (p. 155).

[1] It should be noted that ideally combinatorial systems and ideally iconic systems must take different routes to maximizing distinctiveness. In the former, relatively simple signs can be kept distinct if few in number; in the latter, there are more signs, so greater complexity is required to keep them distinct.

Considering that iconicity is richer and much more frequent in signed languages than in spoken languages (Meier, 2002; Taub, 2001), a possible way to test Goldin-Meadow and McNeill's intuition would be to compare the average levels of combinatoriality in the two modalities. However, there is no clear consensus on how to count meaningless forms in sign language, making such comparisons problematic.

Indeed, with its very high levels of iconicity and very low levels of combinatoriality, ABSL may be the only clear-cut example among natural languages of the tendencies referred to by Goldin-Meadow and McNeill. In this regard, one might consider ABSL to be an exceptional language. However, we claim that what is exceptional in ABSL is not the extreme outcome of the competition between reference and transmission. In fact we consider this the likely outcome in the early stages of any communication system where iconicity is readily available. What is exceptional about ABSL, we would argue, is that this extreme outcome appears to have persisted longer than is typical in other languages. Why this should be the case is not entirely clear. It may have something to do with the nature of the community in which ABSL is used. The Al-Sayyid Bedouin community is a small close-knit one, with relatively low levels of interaction with outsiders. Such communities, termed "esoteric" by Wray and Grace (2007), have been associated with linguistic conservativism and the perpetuation of unusual linguistic features (see Meir, Israel, Sandler, Padden, & Aronoff, 2012 for a comparison of ABSL and Israeli Sign Language in these terms). Nevertheless, while esoteric communities are associated with greater phonological complexity (consistent with what is seen in ABSL), they are also associated with less transparent linguistic forms (Wray & Grace, 2007), which seems at odds with ABSL's high levels of iconicity. It may be that the high level of integration of deaf people into the Al-Sayyid Bedouin community and the consequent use of ABSL by many non-deaf members (Sandler, Meir, Padden, & Aronoff, 2005, p. 2662) play some role here. Because ability among these non-deaf signers is likely to vary more than among deaf signers, it might be that the need for referential efficiency has been prolonged beyond what is typical for other languages. It should be stressed that these explanations remain somewhat speculative, however; indeed, it is possible that ABSL is less exceptional than it appears, and that further research on other newly emerging sign languages will reveal similar patterns of iconicity and combinatoriality.

1.1. The Reference-Before-Transmission Hypothesis

The argument presented above, together with Goldin-Meadow and McNeill's (1999) intuition, leads us to two closely related hypotheses. When emerging communication systems have the capacity for iconicity, the pressure for referential efficiency overcomes the pressure for transmission efficiency, leading to relatively low levels of combinatoriality. On the other hand, when emerging communication systems have little capacity for iconicity, their development will be dominated by the pressure for transmission efficiency, leading to relatively high levels of combinatoriality.

These hypotheses, which we will refer to collectively as the RBTH (for Reference-Before-Transmission Hypothesis), imply three related predictions. The first, which we will refer to as the *Iconic-scaffolding prediction*, is that iconic communication systems will develop faster than non-iconic ones. The second, which we will refer to as the *Combinatoriality prediction*, is that non-iconic communication systems will exhibit higher combinatoriality than iconic systems. The third prediction, which we will refer to as the *Transmission-efficiency prediction*, follows from the second. It is that non-iconic communication systems will exhibit higher transmission efficiency than iconic ones, meaning that they will consist of simpler forms that are easier to produce and perceive.

1.2. Testing the RBTH

One possibility for testing the RBTH might be to compare data from different real-world languages. This is problematic, however. First, as noted above, there is a lack of consensus as to how to count meaningless forms in sign language. And even if there were a consensus, the differences between the phonetics of sign language and the phonetics of spoken language make it unclear how valid a cross-modal comparison between signed and spoken languages could be. These are not the only potential difficulties. Most spoken languages are much older than most sign languages, and – while sign languages generally develop among individuals with little exposure to spoken language – all known sign languages evolved in communities in which spoken languages were well established, meaning that there has been much more potential for speech to influence sign than vice versa. These issues raise serious problems for testing the RBTH using real-world linguistic data.

1.2.1. Testing the RBTH in the laboratory

Testing the RBTH implies several requirements. First, we need to study communication systems developed under conditions that are identical except for the degree to which they afford iconicity. Second, we need to study communication systems that are as independent as possible from pre-established forms of communication and whose degree of residual dependence on them is as similar as possible. Third, we need to study communication systems that developed independently from one another. Finally, we need to study communication systems for which combinatoriality can be measured in the same way.

Recently a line of research has been developed that fulfills these requirements. This line of research, which has been called Experimental Semiotics (Galantucci, 2009), involves studying the emergence and evolution of novel communication systems in the laboratory (see Galantucci, Garrod, & Roberts, 2012, for a review). Work done within Experimental Semiotics has already yielded insights into iconicity and combinatoriality. Studies performed thus far suggest, for example, that – in line with the Iconic Scaffolding Prediction – iconicity aids the bootstrapping of new communication systems (Fay et al., 2013, 2014; Galantucci, 2005; Garrod, Fay, Lee, Oberlander, & MacLeod, 2007; Roberts & Galantucci, 2012; Theisen, Oberlander, & Kirby, 2010). With respect to combinatoriality, experimental studies have examined the role of intergenerational learning (Del Giudice, 2012; Verhoef, 2012; Verhoef, Kirby, & de Boer, 2013) and the role of communication between individuals (Galantucci, Kroos, & Rhodes, 2010; Roberts & Galantucci, 2012). The latter studies, perhaps because iconicity is of particular importance in communicative interaction, have produced preliminary evidence in favor of the RBTH.

1.2.2. Preliminary laboratory evidence in support of the RBTH

Galantucci et al. (2010) studied communication systems developed by pairs of participants playing a videogame. Each turn, players controlled agents that started in two different rooms in a grid and had to find each other. This required coordination, as each player could see only the room their own agent was in. However, the sole means of communication available to players was to make tracings with a stylus on a digitizing pad, which produced real-time signals on both players' screens (Galantucci, 2005). There were two experimental conditions. In both conditions the horizontal location of the stylus on the pad controlled the horizontal location of the signal, but its vertical location was ignored, making it difficult if not impossible to produce conventional graphic forms such as letters or numbers (Fig. 2d). In the *Fast-fading* condition, the signal appeared at a fixed height and disappeared as soon as the stylus was lifted from the pad. In the *Slow-fading* condition, the signal appeared at the top of the screen and scrolled down at a constant speed until, after about 2.5 s, it disappeared from the screen (Fig. 2c). Galantucci et al.'s study is useful for our purposes in two ways. First, as we shall see in Experiments 1 and 2 below, the operationalized distinction between *tracings* (actions performed with the stylus on the pad) and *signals* (visual events derived from the tracings) allows the systematic manipulation of iconicity. Second, Galantucci et al. introduced a numerical procedure to compute an index of combinatoriality through which different communication systems could be soundly compared. This involved identifying the basic forms from which the signals were composed and counting how often they recurred (see Section 4.1.5 for more details, and Fig. 3 for an example).

The results of the study were clear: Fast-fading communication systems were much more combinatorial than Slow-fading ones. On re-examining the signals in Galantucci et al.'s study, furthermore, we found that many of them could be interpreted as iconic, particularly in the Slow-fading condition. In other words, consistent with the RBTH, greater iconicity seems to be associated with lower combinatoriality.

To investigate this association, Roberts and Galantucci (2012) analyzed communication systems developed using a medium similar to Galantucci et al.'s (2010) slow-fading medium. In this study pairs of participants alternated as sender and receiver and played a game in which the receiver had to guess which of a set of animal silhouettes (Fig. 1) the sender was communicating. This resulted in sets of *signs* — pairings of signals and referents. With the

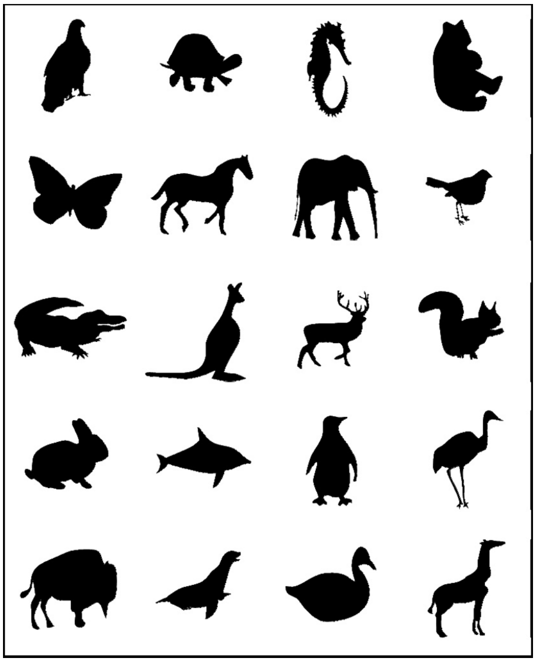


Fig. 1. Animal silhouettes used as referents by Roberts and Galantucci (2012).

communication medium available, it was impossible to create fully iconic signals (i.e., signals that looked exactly like their referents), but it was possible to create partially iconic signals (in which, for example, a jagged line represented teeth or fur). Roberts and Galantucci measured combinatoriality similarly to Galantucci et al. (2010) and measured the signs' *Transparency* by asking naïve judges to match signals to referents. Because iconic signals should be easier to match with their referents, Transparency can be seen as a proxy for iconicity. The results were again consistent with the RBTH: Combinatoriality and Transparency were negatively correlated.

These two studies provide encouraging data, but are not sound tests of the RBTH for two reasons. First, iconicity was not manipulated in either study. Second, neither study allowed combinatorial structure to be clearly distinguished from compositional structure. In Experiment 1, described below, we manipulate the opportunity for iconicity, minimizing it in one condition and maximizing it in the other, while at the same time minimizing the opportunity for compositionality in both conditions. This allows us to directly test the predictions of the RBTH.

2. Experiment 1

2.1. Method

2.1.1. Participants

40 students from universities in New York City participated in pairs (henceforth *dyads*) for money or course credit. Students with deficits in color vision, or who had participated in similar studies before, were excluded from participating.

2.1.2. Materials

Participants played a guessing game using two linked computers running custom-designed software. They communicated with each other using *Wacom Bamboo Pen* CTL-460 digitizing pads and styluses.

2.1.3. Guessing game

Two participants (henceforth *players*) sat at computers in separate locations and played a cooperative guessing game similar to that of Roberts and Galantucci (2012). Each player saw the same set of ten items (henceforth *referents*) displayed in different random locations in a 5-by-5 on-screen grid (Fig. 2a and b). Players took turns as *sender* and *receiver*. At the start of a turn, one referent was chosen as the *target* and highlighted in the center of the sender's screen. The sender's goal was to communicate the identity of the target to the receiver; the receiver's goal was to select this referent. As in the studies by Galantucci et al. (2010) and Roberts and Galantucci (2012), the players were in separate rooms, and the sender could communicate only by using a stylus to make tracings on a digitizing pad (sampled at approximately 66 Hz). These tracings produced a real-time signal in a communication panel on the top-right corner of both players' screens. Tracings were systematically transformed into signals such that players could not effectively use pre-established forms of communication, but had to cooperatively develop a novel communication system from scratch. Specifically, the horizontal component of the tracing determined the horizontal component of the signal but the vertical component was replaced with upward movement at a constant rate, as if the sender was drawing on scrolling paper. (See Fig. 2c and d for examples.) The stylus could be lifted off the pad to create a gap in the signal, but variation in pressure made no difference. Participants were made aware before the game started that both sender and receiver would see the signal as it was being produced. In fact, observing one's own signal as sender was the sole means of coming to understand the relationship between tracing and signal; at no point were participants told explicitly how the tracing-to-signal transformation worked, and there were no demonstrations or practice rounds before the game started. Both players did however receive feedback on their communicative success. A line graph on the lower right of the screen indicated what proportion of referents the dyad had mastered (that is, had successfully communicated at least 75% of the time). This graph was updated after each turn. Additionally, for three seconds at the end of each turn the receiver saw what the target referent was, and the sender saw which referent (if any) the receiver had selected.

All 10 referents were visible to both players throughout, but their grid locations were changed at random every turn and differed for the two players, making it impossible to communicate the target by referring to its location. In early turns, to ease players into the game, a focus set of four referents (Fig. 2b) were presented as targets twice each in every twelve turns, with the target in the remaining four

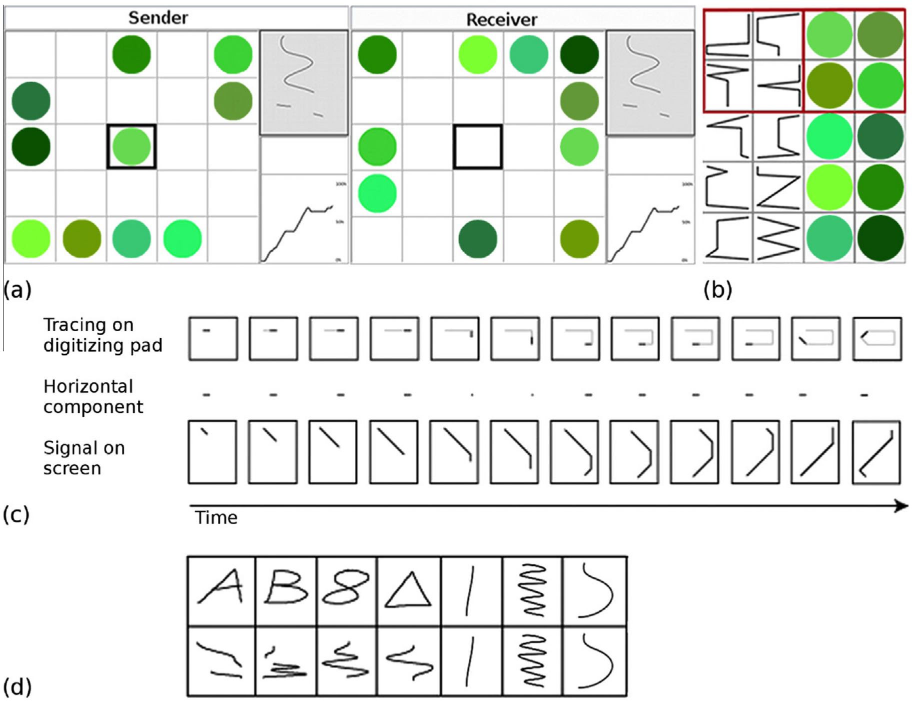


Fig. 2. (a) Example screens. For the sender, the black box highlighted the target referent (always in the center); for the receiver, the black box could be moved to highlight the referent the receiver believed was the target. The sender's signal appeared for both players in the communication panel on the top right, while the bottom-right area contained a plot of the dyad's score over time. (b) Line and Color referents. Focus sets are highlighted in red. (c) Transformation of tracings into signals; (d) examples of how different tracings would appear as signals.

turns chosen randomly from the other six referents (the order in which particular targets appeared within this twelve-turn cycle was randomized). If a dyad's success-rate reached 75% for each focus-set member, the bias disappeared and all referents appeared thenceforth with equal probability. The game could end in one of two ways: (a) if a dyad reached at least 75% success on every referent; (b) after 90 min, if (a) had not occurred.

2.1.4. Manipulation: opportunity for iconicity

The experiment had a between-subjects design with two conditions, which differed in the opportunity for iconicity. A signal is iconic if intuitively motivated by its referent. We manipulated the opportunity for iconicity by varying the type of referent players had to communicate.

In the *Mimable* condition, each referent consisted of a continuous line with four straight segments (Fig. 2b). These *line referents* differed from each other with respect to geometric profile only and were designed so that they could be clearly *mimed* using the communication medium available (that is, signals could be created that strongly resembled these referents and were thus highly iconic). In the *Non-mimable* condition, each referent consisted of a circle colored a different shade of green (Fig. 2b). These *color referents* — which were all the same size and shape and differed from each other only with respect to color — were designed so that they could not be mimed with the communication medium available.

Apart from being chosen to differ in terms of mimability, both sets of referents were designed to reduce as far as possible the opportunity for compositionality (the recombination of meaningful forms) as a communicative strategy. This meant that the reuse of basic forms in participants' communication systems could be identified as genuinely combinatorial.

So that combinatoriality could be measured in the same way in both conditions, the communication medium was identical for all participants. Several points should be noted regarding this methodological choice. First, because the medium was identical, there was nothing to prevent players from developing combinatorial (or non-combinatorial) signals in either condition. Second, although the *opportunity* for iconicity varied between conditions, it was equally possible to create non-iconic signals for either of the two referent sets. For example, players could adopt the basic forms of Morse code (dots and dashes were easy to produce with the game's communication device) and use them to develop a simple and robust communication system in both conditions. Finally, it was possible in the Mimable condition to create signals that had elements of both iconicity and combinatoriality (an example of such a signal can be seen in Fig. 5).

Our expectation that signals in the Mimable condition would be iconic and non-combinatorial hinges on the hypothesis that the pressure for referential efficiency trumps the pressure for transmission efficiency, leading participants to produce highly iconic signals at the expense of greater complexity. On the other hand, we expected that the difficulty in satisfying the pressure for referential efficiency in the Non-mimable condition would give transmission efficiency the upper hand, leading participants to produce signals composed of a small set of relatively simple meaningless forms (rather than creating a unique form for each referent, as was predicted for the Mimable condition).

2.1.4.1. Manipulation check. To check that our manipulation worked as intended, we measured the Transparency of the sign-sets in the same way as Roberts and Galantucci (2012). For each condition four naïve judges matched referents with the signals that referred to them. Judges first gained an understanding of the game by playing a few turns as both sender and receiver. They then saw a display containing one player's signals (as playable videos) along with the referents they referred to (both displayed in a random order). Each judge was given as much time as needed to match the former with the latter; then another player's sign-set would appear. Each judge evaluated one sign-set from every dyad, and every sign-set was shown to two judges. (One dyad's sign-sets were excluded as they contained containing only one sign.) For every sign-set the mean number of correct matches was calculated and divided by the number of signs in the set to produce an index from 0 to 1, where 0 corresponds to a non-transparent sign-set and 1 to a maximally transparent set. A Monte Carlo test was then performed by randomly rematching signals with referents 100,000 times and comparing these matches with the judges' matches.

The results suggested that participants in the Mimable condition were creating highly iconic signals (Mean Transparency = 0.995; $p < 0.001$); in fact only one of the four judges made an incorrect match, confusing two signals in one sign set. In the Non-mimable condition, by contrast, judges were not able to match signals with referents at above chance level ($M = 0.29$; $p = 0.46$), suggesting that iconicity had been successfully blocked in this condition.[2]

2.1.5. Measures

To test the predictions of the RBTH we used a number of measures, which will be described in what follows.

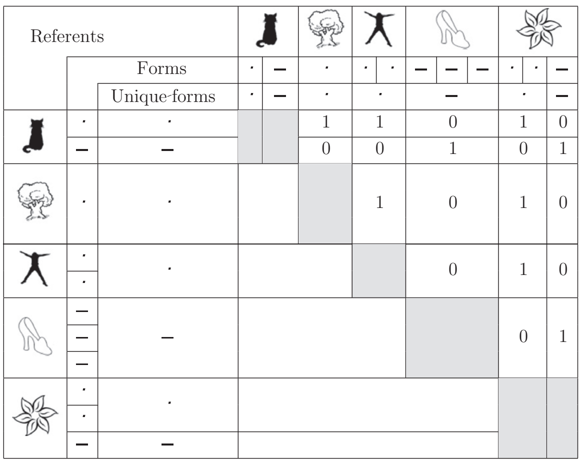
2.1.5.1. Set-size and initial delay: Testing the iconic-scaffolding prediction. According to the Iconic-scaffolding prediction, iconic communication systems should develop faster than non-iconic ones. This means that players in the Mimable condition should establish more signs by the end of the game than players in the Non-mimable condition. To test this, a sign-set was constructed for each player following the game. This consisted of every referent on which the dyad reached at least 75% success, paired with the last successful signal the player used to communicate it. The size of the sign-set (henceforth *Set-size*) provided a measure of communicative success: The larger the set, the more meanings players could communicate.

[2] It is important to note that Transparency is not equivalent to iconicity, but is a proxy for it; it is in principle possible that certain participants' signals in the Non-mimable condition were iconic in a way that was obvious to them, but not to others. However, we feel that the Transparency results, taken together with the inherent difficulty in producing iconic signals in this condition, allow us to be confident that the opportunity for iconicity was minimized in this condition.

Set-size was initially calculated for each player; then the mean value was calculated for each dyad.

Iconicity may also give senders a head-start in producing signals. We therefore measured how long it took them at the start of each turn before touching the stylus to the pad (*Initial delay*). Unlike Set-size, this measure was likely to be especially relevant to early stages of the game, when participants were still establishing signals. We therefore divided each game into three parts (the *first third*, the *second third*, and the *final third*) and calculated the mean Initial delay for each part. In all cases the Initial delay was first calculated for each player and then the mean value was calculated for each dyad.

2.1.5.2. Combinatoriality: Testing the Combinatoriality prediction. The Combinatoriality prediction states that non-iconic communication systems will exhibit higher combinatoriality than iconic systems. We measured *Combinatoriality* using a slightly modified version of the Form Recombination Index (FRI) developed by Galantucci et al. (2010). The logic of the FRI is as follows. A highly combinatorial system, like Morse code, is one in which a small number of basic units recur across a large number of signals. To measure Combinatoriality, therefore, one must identify the basic units of the communication system in question and count how often they recur across signals. The FRI is calculated simply as the number of recurrences divided by the number of possible recurrences (Fig. 3). For example, if the signals of a communication system were composed entirely of dots (one dot for "cat" and two dots for "tree", say), then the system's FRI would be 1, meaning that it was perfectly combinatorial, as its one unit recurred everywhere it possibly could. Introducing extra forms reduces the FRI value; see Fig. 3 for an example of a small Morse-code-like system in which signals are composed of both dots and dashes (a dot and a dash for "cat", a dot for "tree" etc.) and for which the FRI is 0.47.



Referents											
	Forms	·	–	·	·	·	– – –			· ·	–
	Unique-forms	·	–	·	·		–			·	–
	· / ·			1	1		0			1	0
	– / –			0	0		1			0	1
	· / ·				1		0			1	0
	· · / ·						0			1	0
	– – – / –									0	1
	· · – / · –										

Fig. 3. An example of the procedure used for computing the Form Recombination Index (FRI). The procedure is applied to a small communication system in which signals composed of two basic forms (dots and dashes) are used to communicate about five referents (a cat, a tree, a person, a shoe, and a flower). First, repetitions of the same form in the same signal are ignored, leaving a set of unique forms. The recurrence of a unique form between signals is indicated by a "1" in the table. There are 9 such recurrences, and 10 non-recurrences, yielding an FRI of 0.47 [9/(10 + 9)].

Both steps of the FRI — identifying units and counting how often they recur — are less trivial than they might appear from our examples. The first step is non-trivial because, in principle, there are a great number of ways in which signals in our experiment could be segmented into units. (Similar problems, it should be noted, are faced in the analysis of real-world languages.) However, observation of and feedback from participants indicated that an approach with particularly strong psychological reality was to segment signals into *forms*—shapes in the signal created by contact between the stylus and the pad, and separated by gaps when the stylus is raised off the pad (see Fig. 5 for example signals). We treated these forms as the basic units of the players' communication systems.

The challenge of the second step — identifying recurrences of forms — is to identify whether two forms in different signals should be considered equivalent (and, therefore, as recurrences of each other). To avoid calculating equivalence subjectively, we used an algorithm for comparing forms based on that used by Galantucci et al. (2010). The first step in this algorithm was to compute the mean value of each form (i.e., the mean horizontal position of the stylus) and determine how often the form crossed this value (henceforth *mean-crossings*). Next, the proportion of the form that fell between each mean-crossing (henceforth *form proportions*; Fig. 4) was computed; this meant that the shape of the form could be expressed as an ordered series of form proportions. Any two forms were treated as equivalent if they fulfilled two criteria: (a) they had the same number of mean-crossings, and (b) the respective form proportions had values that were within twenty percent of each other. We used this equivalence algorithm twice. First we compared forms within a signal. If two or more within-signal forms were equivalent, one was randomly selected to represent the others, reducing the set of forms to a set of *unique* forms (in this respect our measure differs from that used by Galantucci et al., 2010). Then the equivalence test was used to compare unique forms across signals.

Once form equivalences had been computed, we counted how often the same form recurred across signals and divided the number of recurrences by the number of recurrences plus the number of non-recurrences (Fig. 3) to produce an index ranging from 0 and 1 (where 0 corresponds to a complete absence of combinatoriality and 1 corresponds to maximal combinatoriality). Combinatoriality was initially calculated for each player; then the mean value was calculated for each dyad. Because identifying unique forms involved randomly selecting one form to represent others, the Combinatoriality index varied slightly each time it was calculated. We therefore calculated it 1000 times and used its mean value for our analyses.

2.1.5.3. Complexity, distinctiveness, and production time: Testing the transmission efficiency prediction. The transmission-efficiency prediction states that non-iconic

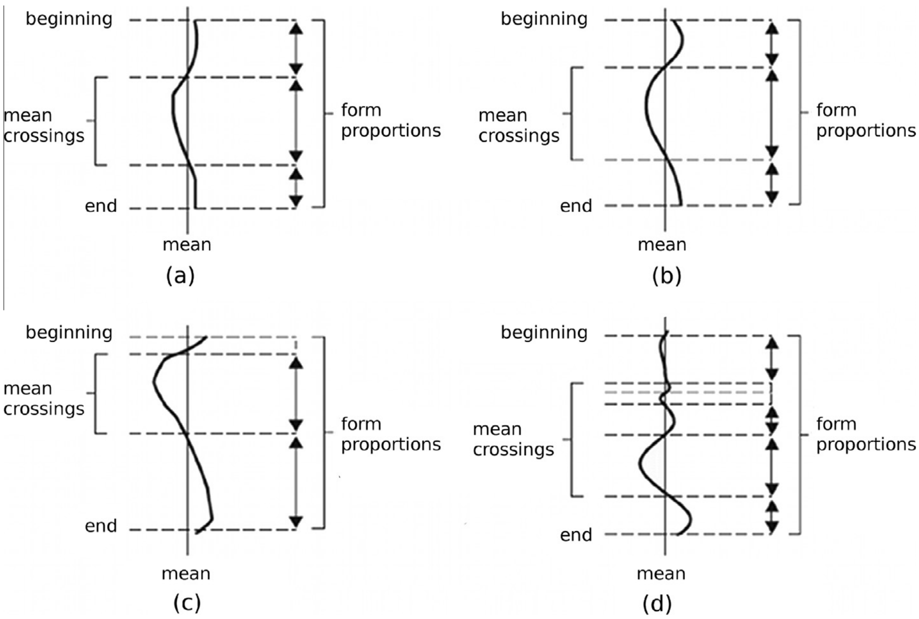


Fig. 4. Form equivalence test. Forms in (a) and (b) are equivalent. The forms in (a) and (c) are not, because form proportions are not within 20% of one another. The same is true of the forms in (b) and (c). The form in (d) is not equivalent to any of the other forms because it has a different number of mean crossings (adapted from Galantucci et al., 2010). *Note*: Although these figures show forms as graphical signals, the equivalence test was in fact calculated based purely on the horizontal position of the stylus on the pad; the same test could therefore be employed for signs in Experiment 2, in which signals looked very different.

communication systems will exhibit higher transmission efficiency than iconic ones. There are two ways in which this might be the case: Signals might be simpler to produce or they might be easier to distinguish perceptually.

To ascertain how simple a given signal was to produce we first broke it into forms, as for the FRI, and measured the Complexity of each form by calculating the entropy of its local curvatures (see Page, Koschan, Sukumar, Abidi, & Abidi, 2003, for a detailed description of this measure). The basic logic of this measure is that, when forms are more complex, the distribution of their local curvatures is highly varied, leading to higher overall entropy. Since tracing forms with highly varied curvatures is harder than tracing forms with little variation in curvature, we used form complexity as a proxy for ease of production (the lower the complexity, the easier the production). Because this measure concerned production, not perception, it was calculated based on the tracing, not the on-screen signal. That is, it took into account the actual vertical coordinates of the participant's stylus on the digitizing pad. Once the Complexity of each form had been calculated, these measures were used to calculate the mean Complexity for each tracing.

Because more complex signals are likely to take longer to produce, we also calculated how long it took senders to produce signals (*Production time*). This was measured from the moment that the stylus first touched the pad to the last time it was raised from the pad during the turn.

To ascertain the perceptual efficiency of sign sets we measured their *Distinctiveness* by asking 24 naïve judges to distinguish between them. Each judge saw a video of every signal from one member of each dyad, presented in a random order.[3] After a signal had been displayed, the screen turned black for one second, and another signal was displayed. Half the time this second signal was the same as the first; half the time it was a different signal selected randomly from the same sign-set. The judge's task was to decide if the two were the same or different. Finally, the number of correct guesses was divided by the number of total guesses to produce an index from 0 to 1, where 0.5 meant that signals could not be distinguished at all, and 1 meant that signals were maximally distinct. (A value of less than 0.5 would have suggested that participants were performing at below chance level.)

2.2. Results

Example signs can be seen in Fig. 5. The measures described above were applied to the data and the results of this analysis are displayed in Fig. 6.

2.2.1. Set-size and Initial delay

As predicted by the RBTH, Set-sizes were smaller in the Non-mimable condition ($M = 4.6$, $SD = 2.99$) than in the Mimable condition ($M = 9.75$, $SD = 0.26$), $p < 0.001$[4] (mean

[3] Note that this measure was performed after Experiment 2 (see below), so judges also saw signals from that study.

[4] All p values reported in this paper were derived using a two-tailed Monte Carlo test, in which the data was resampled 100,000 times without replacement.

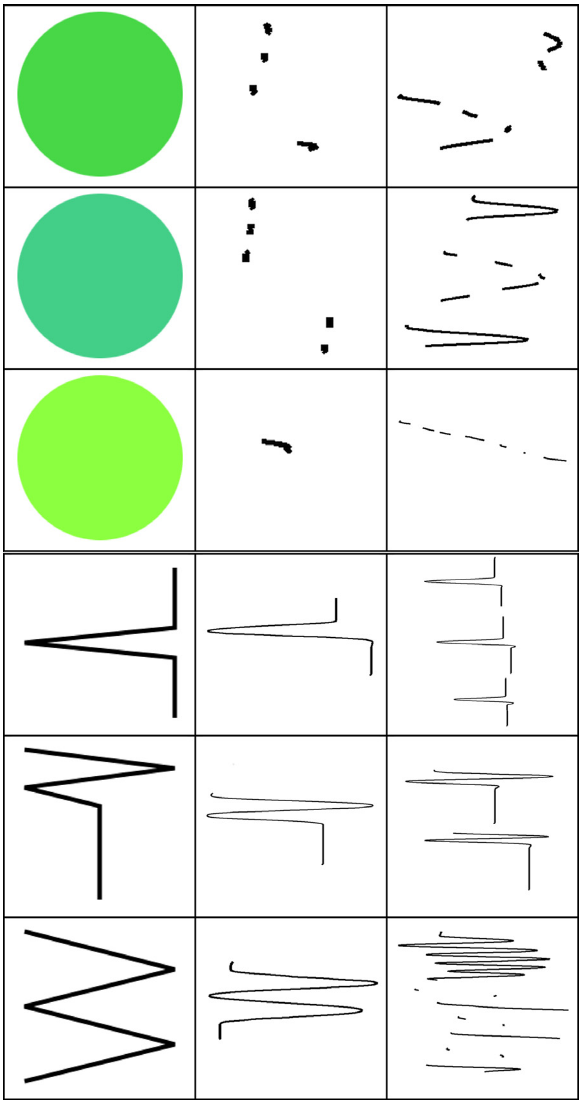


Fig. 5. Example signs from different dyads in both conditions. (*Note*: smaller signals have been enlarged for clarity.)

difference 5.15), suggesting that participants in this condition found it harder to construct sign-sets. Indeed, no dyad in the Non-mimable condition reached a 75% success rate on all ten referents, while every dyad in the Mimable condition did so.

The Initial delay before touching stylus to pad was significantly longer in the Non-mimable condition ($M = 2.8$ s, $SD = 0.8$) than in the Mimable condition ($M = 1.57$ s, $SD = 0.45$; mean difference 1.23 s; $p < 0.001$). The difference was smallest in the first third of the game, when players in both conditions were less practiced (mean difference 1.03 s, $p = 0.001$). In both conditions the delay grew shorter over the course of the game, but the gap between conditions was at its largest in the second third (second third 1.37 s, $p < 0.001$; final third 1.28 s, $p < 0.001$).

2.2.2. Combinatoriality

As predicted by the RBTH, Combinatoriality was higher in the Non-mimable condition ($M = 0.25$, $SD = 0.3$) than in the Mimable condition ($M = 0.03$, $SD = 0.02$), $p < 0.01$ (mean difference 0.22). Because Set-size varied between conditions, we reran the original Combinatoriality analysis with Set-size as a covariate. The pattern of results did not change: $F(1,16) = 6.17$, $p < 0.05$.

2.2.3. Complexity, production time, and distinctiveness

As predicted by the RBTH, tracings in the Non-mimable condition ($M = 0.29$, $SD = 0.3$) were less complex than those in the Mimable condition ($M = 0.97$, $SD = 0.09$), $p < 0.001$ (mean difference 0.68). The pattern of results did not change with Set-size as a covariate $F(1,17) = 50.79$, $p < 0.001$.

Production time for signals was shorter in the Non-mimable condition ($M = 3.86$ s, $SD = 1.48$) than in the Mimable condition ($M = 4.65$ s, $SD = 1.75$; mean difference 0.79 s), but the difference was not significant ($p = 0.29$). In both conditions, participants spent longer producing their signals as the game went on. In the Non-mimable condition mean Production time increased from 3.58 s in the first third to 3.99 s in the final two thirds of the game. In the Mimable condition it increased from 3.89 s in the first third to 4.93 s in the second third to 5.11 s in the final third. In no case was the difference between conditions significant.

While signals in the Non-mimable condition ($M = 0.89$, $SD = 0.08$) appear to have been slightly easier to distinguish than those in the Mimable condition ($M = 0.85$, $SD = 0.11$), the difference was not significant (mean difference 0.04, $p = 0.16$).

2.2.4. Discussion

The results of Experiment 1 supported the RBTH. In line with the Combinatoriality prediction, combinatorial structure emerged when iconicity was unavailable as a strategy, but only to a minimal extent when it was available. This result is not due to a difference in the *availability* of combinatoriality as a strategy; rather, as our manipulation check shows, it is due to participants' overwhelming rejection of this strategy in favor of iconicity. This is because, in line with the Iconic-scaffolding prediction, iconicity made it easier to establish new signs, an advantage that increased over the course of the game, as reflected in the results of the Initial delay measure. This advantage, however, came at a cost. Consistent with the Transmission-efficiency prediction, the non-combinatorial iconic signs were more complex to produce than the non-iconic combinatorial ones.

It remains possible, however, that the differences we found between the two conditions were the result of differences not in the opportunity for iconicity (that is, in the relationship between the signaling system and the referents) but rather in some property of the referents themselves. For example, it might be that one set of referents could be discriminated or remembered more easily, or that the Non-mimable referent set afforded compositional sign systems, which would look superficially combinatorial.

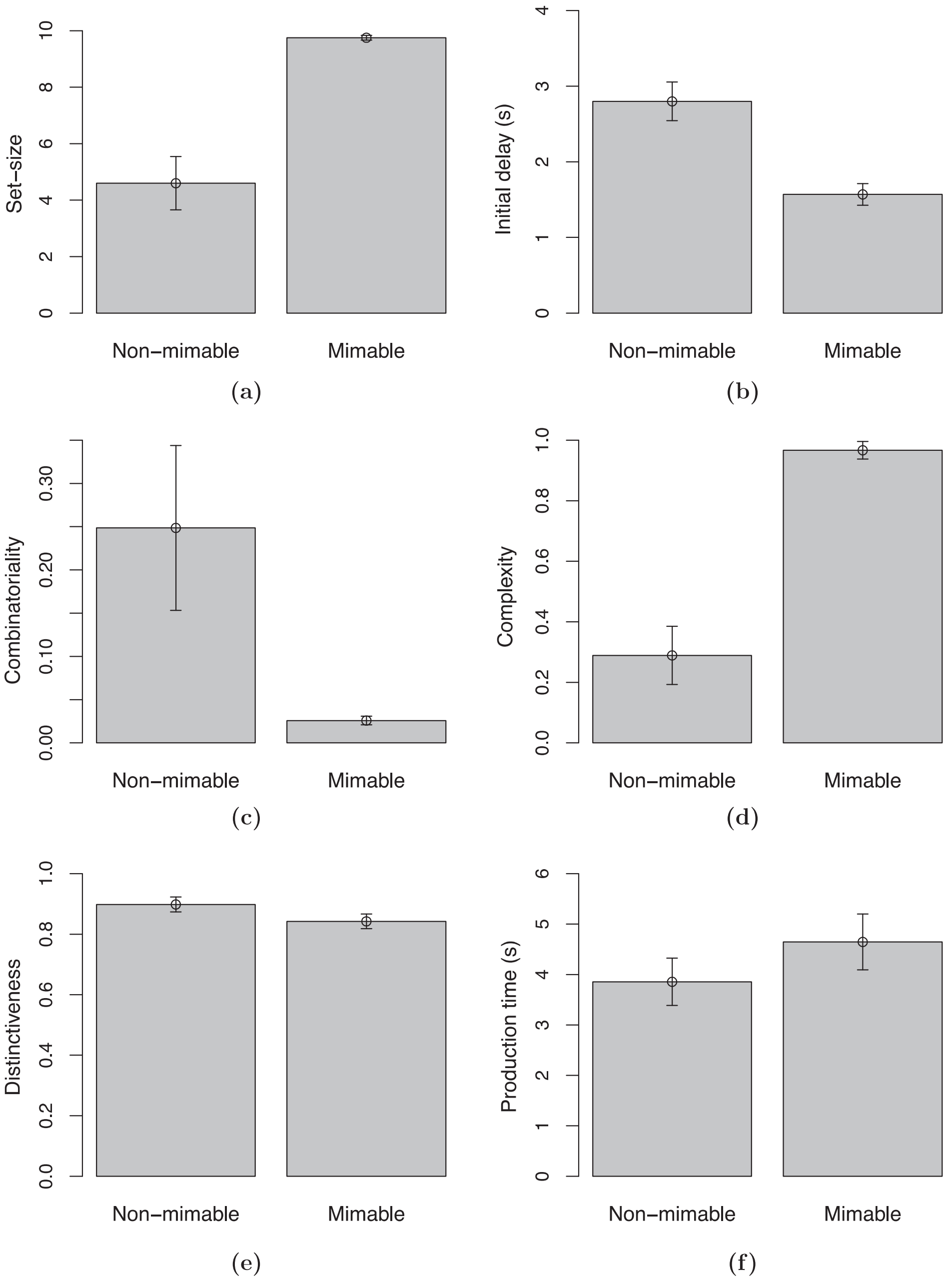


Fig. 6. (a) Mean Set-size ($N = 10$ for both conditions); (b) Mean Initial delay ($N = 10$); (c) Mean Combinatoriality ($N = 10$ for Mimable condition; $N = 9$ in Non-mimable condition, where Set-size was too small to calculate Combinatoriality for one dyad); (d) Mean Complexity ($N = 10$); (e) Mean Distinctiveness ($N = 10$); and (f) Mean Production time ($N = 10$) Error bars indicate standard error.

3. Experiment 2

To eliminate the possibility that the results of Experiment 1 were due to some property of the referents, we carried out a second experiment using the same game, referents, stylus and pad as in Experiment 1, but in which the tracing-to-signal transformation was changed to allow the color referents to be mimable.

3.1. Method

3.1.1. Participants and materials

20 students from universities in New York City participated in dyads for money or course credit. Students with deficits in color vision, or who had participated in similar studies before, were excluded from participating. The same materials were used as in the first experiment.

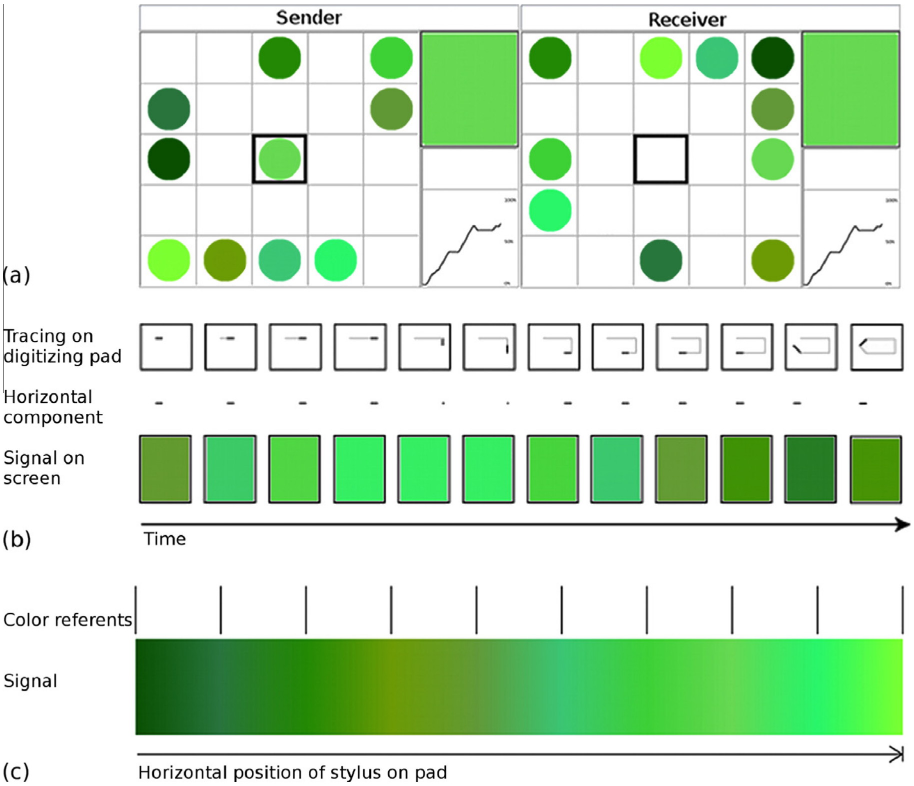


Fig. 7. (a) Example player screens in Experiment 2. (b) Transformation of tracings into signals. (c) Color range onto which the horizontal position of the stylus was mapped to produce Color signals; vertical lines indicate colors used for Color referents.

3.1.2. Guessing game

The guessing game was the same as in the first experiment except for the tracing-to-signal transformation, which was modified to allow the color referents to be represented iconically. The communication panel appeared gray whenever the stylus was not touching the pad. Whenever the stylus was touching the pad, its horizontal position determined the color of the entire communication panel; vertical position was ignored (Fig. 7a and b). More specifically, the horizontal position of the stylus was mapped onto a range of greens such that the 10 Color referents (Fig. 2b) corresponded to 10 equidistant points on the pad (Fig. 7c). For example, if the stylus were touching the leftmost edge of the pad, the communication panel would be the same color as the darkest of the ten referents. If the stylus were touching the rightmost edge, the panel would be a lighter color. If the stylus were moved horizontally along the pad, the panel would change color as the stylus moved (Fig. 7c). As in the first experiment, participants were not told explicitly how tracings related to signals, but both participants could observe their communication panels change color in real time.

3.1.3. Measuring combinatoriality

In Experiment 1 strokes of the stylus on the pad produced shapes on the communication panel separated (when the stylus was raised) by gaps (Fig. 2a and b; see also example signals in Fig. 5). In the Mimable-color condition, the same strokes of the stylus produced periods of color in the communication panel separated (when the stylus was raised) by periods of gray. In spite of this difference, the stylus could be used in precisely the same way as in Experiment 1 (for example, players could easily develop a Morse-code-like system as described above), and forms could be defined in the same way as visual units, caused by contact between the stylus and the pad and separated by gray space. We therefore measured Combinatoriality in the Mimable-color condition as in Experiment 1: We identified forms as periods in the signal in which the stylus was in contact with the pad and counted how often the same form recurred across signals. Because the form equivalence test (Fig. 4) was based entirely on horizontal stylus-position, not on how signals appeared onscreen, it could be used for Experiment 2 even though the signals looked very different from those in Experiment 1.

One change was necessary, however. In Experiment 1 the relative horizontal position of two forms on the pad was ignored for the purposes of determining if one was a recurrence of the other; in Experiment 2, this was clearly inappropriate, as the identity of a form in this condition depended on the horizontal position of the stylus on the pad. Therefore, one form was considered a recurrence of another only if the horizontal distance between the two was no greater than 5% of the width of the communication panel. In the interest of a fair comparison between this condition and the conditions of Experiment 1 — in which participants might or might not have distinguished forms based on their position — the combinatoriality indices for

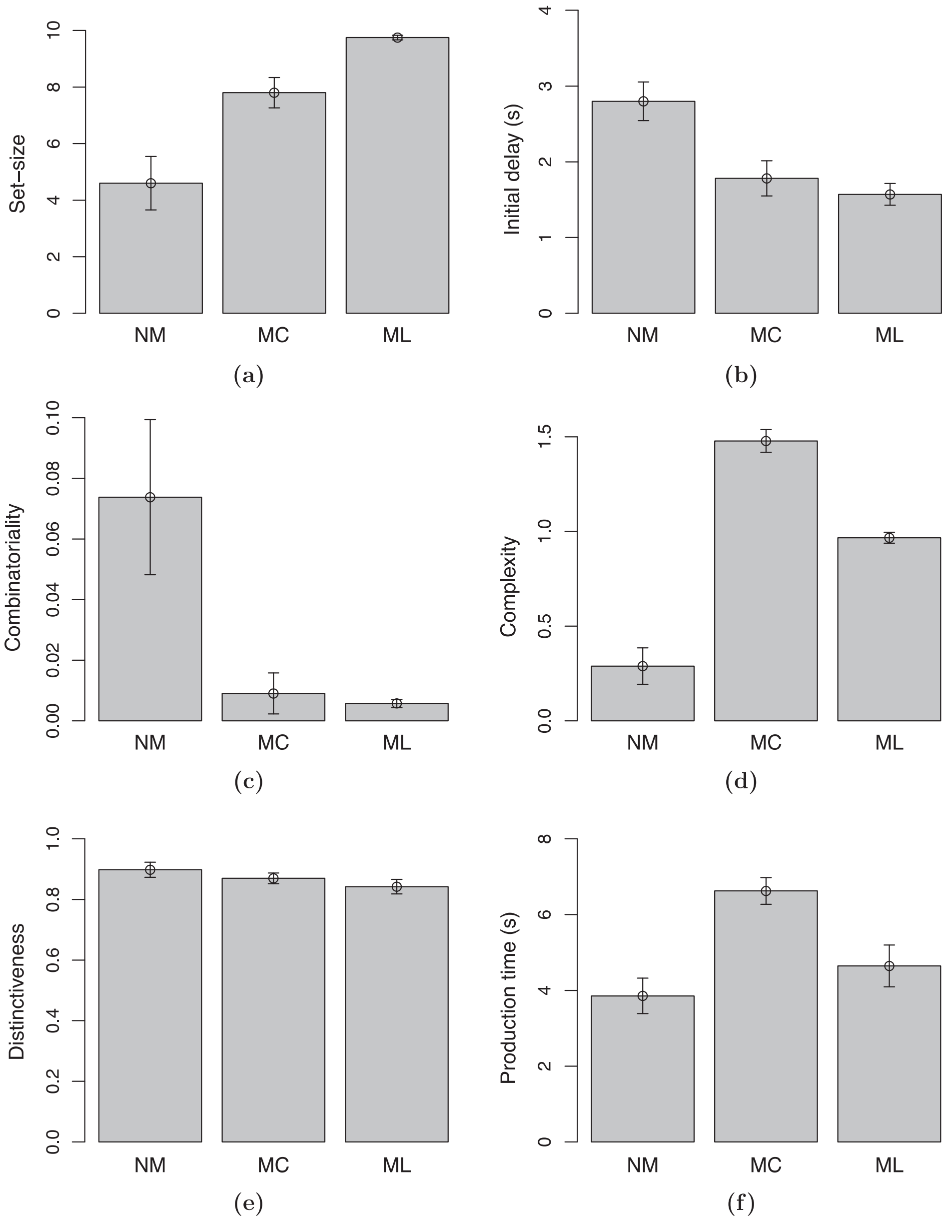


Fig. 8. Results for all conditions. (Experiments 1 and 2). (a) Mean Set-size; (b) Mean Initial delay; (c) Mean Combinatoriality (incorporating horizontal-position constraint); (d) Mean Complexity; (e) Mean Distinctiveness; and (f) Mean Production time. NM = Non-mimable; MC = Mimable-color; ML = Mimable-line. *N* = 10 for MC. Error bars indicate standard error.

the Mimable-line and Non-mimable conditions were recalculated with the same horizontal-position constraint.[5]

3.2. Results

It became apparent through piloting that, in spite of the new tracing-to-signal transformation, participants could still mime the Line referents by tracing their shape on the pad (even though this did not produce an iconic signal) and, indeed, consistently pursued this option.[6] Because this thwarted the iconicity manipulation, only the condition with Color referents was run. The results of this condition (henceforth the *Mimable-color* condition) were then compared with those of the Non-mimable and the Mimable (henceforth *Mimable-line*) conditions from Experiment 1. Results are presented in Fig. 8.

[5] This reduced mean Combinatoriality in the Non-mimable condition from 0.25 to 0.07, and in the Mimable condition from 0.02 to 0.006. The difference between conditions remained, however: mean difference 0.07, $p < 0.01$.

[6] This is in fact consistent with the findings of Experiment 1, and with the RBTH, inasmuch as it implies an attraction to iconicity. However, it is interesting to note that, in this case, iconicity involves highly iconic *tracings* that produce *signals* that are iconic only in an extremely indirect way. The receiver would of course be exposed only to the latter.

3.2.1. Manipulation check

The signals' Transparency was measured in the same way as in Experiment 1. As in the Mimable-line condition of Experiment 1, Transparency was high ($M = 0.82$; $p < 0.001$), suggesting that participants were creating highly iconic signals.

3.2.2. Set-size and Initial delay

Set-sizes were smaller in the Mimable-color condition ($M = 7.8$, $SD = 1.7$) than in the Mimable-line condition, $p < 0.01$ (mean difference 1.95), but greater than in the Non-mimable condition: $t(14.29) = 2.94$, $p = 0.01$ (mean difference 3.2). This likely indicates that, while players found the signaling medium less intuitive in the Mimable-color condition, the opportunity it gave them for iconicity still helped them in constructing a communication system.

Initial delay in the Mimable-color condition ($M = 1.78$ s, 0.73) was significantly shorter than in the Non-mimable condition (mean difference 1.02, $p < 0.01$); the difference was at its largest in the second third of the game (1.38 s) and at its smallest in the final third (0.84 s).

Overall initial delay was longer than in the Mimable-line condition (overall mean difference 0.2 s), but the difference was significant only in the final third of the game (0.44 s, $p = 0.049$).

3.2.3. Combinatoriality

The results were consistent with those from Experiment 1. Combinatoriality in the Mimable-color condition ($M = 0.009$, $SD = 0.02$) was lower than in the Non-mimable condition ($M = 0.07$, $SD = 0.08$): $p = 0.01$ (mean difference 0.06), and not significantly different from Combinatoriality in the Mimable-line condition ($M = 0.006$, $SD = 0.004$): $p = 0.95$.

3.2.4. Complexity, production time, and distinctiveness

The transmission efficiency of signs was measured as in Experiment 1. Complexity was higher in the Mimable-color condition ($M = 1.48$, $SD = 0.19$) than in both the Non-mimable condition, $p < 0.001$ (mean difference 1.06), and the Mimable-line condition, $p < 0.001$ (mean difference 0.51).

Production time was significantly longer in the Mimable-color condition ($M = 6.62$ s, $SD = 1.12$) than in both the Non-mimable condition (mean difference 2.77 s, $p < 0.001$) and the Mimable-line condition (mean difference 1.98 s, $p < 0.01$). As in Experiment 1 mean Production time increased between the first and second thirds of the game (from 6.39 s to 6.9 s). However, it then decreased in the final third (to 6.58 s). Only in the first and second thirds of the game was Production time in the Mimable-color condition significantly different from Production time in the Mimable-line condition (mean differences 2.5 s and 1.96 s; $p < 0.01$). In all three thirds, there was a significant difference between Production time in the Mimable-color and Non-mimable conditions (mean differences 2.8 s, 2.9 s, and 2.6 s respectively; $p < 0.001$).

Distinctiveness was lower in the Mimable-color condition ($M = 0.85$, $SD = 0.1$) than in the Non-mimable condition (mean difference 0.04) and higher than in the Mimable-line condition (mean difference 0.001), but the difference was not significant in either case ($p = 0.4$ and 0.37 respectively).

3.3. Discussion

The results of Experiment 2 suggest that the difference in Combinatoriality in Experiment 1 was due to the difference in the opportunity for iconicity, not to any property of the referents themselves. Once participants were given the opportunity to represent the Color referents iconically, the pressure for referential efficiency trumped the pressure for transmission efficiency, as it did in Experiment 1. This occurred despite signals in the Mimable-color condition being even less efficient to produce than signals in the Mimable-line condition.

4. General discussion

Our results clearly supported the RBTH: We successfully manipulated iconicity in novel communication systems and observed differences in their degree of combinatoriality consistent with the hypothesis. There is now a compelling case to be made for Goldin-Meadow and McNeill's (1999) intuition that combinatoriality emerges where the opportunity for iconicity is reduced.

It is hard not to be struck by the extent to which iconicity was adopted as a strategy in the Mimable conditions (and the extent to which combinatoriality was *not* adopted, even though, as can be seen in Fig. 5, they could potentially have co-occurred). This is particularly notable when one considers that the vast majority of known languages exhibit extensive combinatoriality, and rather less iconicity.[7] In this context it is important to note that nothing in any condition of our study prevented participants from creating non-iconic signals. Indeed, as our efficiency measures revealed, there was a pressure against iconic signals, which were more complex and thus harder to produce. However, as revealed by the differences in Set-size and Initial delay between the Non-mimable and Mimable conditions, this cost was offset by the fact that iconicity greatly assisted players with the process of establishing signs (cf. Fay et al., 2013; Perniss et al., 2010). In doing so it hindered the emergence of simpler units, leading to low levels of combinatoriality. We believe that this core mechanism, exposed in our experiments, is the mechanism that delayed the emergence of combinatoriality in ABSL.

Two limitations should be noted, both deriving from the nature of the referents used in this study. The first is that, as described above, both sets of referents were designed to eliminate the opportunity to create compositional sign systems. This is because the study was designed to focus on the emergence of combinatoriality alone (which we define as the recombination of meaningless units), whose relationship with iconicity is likely to be different from the relationship between iconicity and compositionality.

[7] Though recent research suggests that the extent of iconicity in language has been significantly underestimated (Monaghan, Shillcock, Christiansen, & Kirby, 2014; Perniss et al., 2010).

Had compositional structure (the recombination of meaningful units) emerged to a significant extent, it could have led to spuriously high measures of combinatoriality. We consider this methodological choice to be justified on both theoretical and experimental grounds, but we note that the reality of language emergence is likely to be more complex. While our results suggest that iconicity delays the emergence of combinatoriality, we would not want to extend that claim to the emergence of compositionality; nor would we claim that compositionality itself delays the emergence of combinatoriality. Indeed, the two kinds of structure may often emerge together in real-world communication systems (e.g., Goldin-Meadow, Mylander, & Butcher, 1995; Senghas, Özyürek, & Goldin-Meadow, 2010). Furthermore, distinguishing between genuinely meaningful and genuinely meaningless units may not always be straightforward (see Roberts & Galantucci, 2012, pp. 310–312, for a discussion of the difficulties involved; for discussions of similar problems in a non-experimental context see Stokoe, 1991, and Ladd, 2012). Likely for this reason, previous Experimental-semiotic research has tended either to focus on one or the other kind of structure or to conflate the two (for instance, Del Giudice, 2012, investigated the emergence of "sub-lexical structure").

The second limitation concerns the extent to which the Mimable referents afforded iconicity. In both the Mimable-line condition and the Mimable-color condition it was possible to create signals that were close to isomorphic with what they referred to, a feature participants overwhelmingly took advantage of. The reason for this choice was to distinguish the conditions as clearly as possible. However, it has two consequences. First, it means that this study is limited in what light it can shed on the emergence of combinatoriality where the opportunity for iconicity is more constrained. To some extent that question was addressed by Roberts and Galantucci's (2012) study, in which a single set of referents was used for which full iconicity was not possible (Fig. 1). In that study Combinatoriality was negatively correlated with Transparency. However, we feel that this is an area of study that is very much open for further exploration.

A second consequence of such complete iconicity is that the apparent level of combinatoriality for the signals in question is necessarily the same as for the referents themselves. This means that a fully iconic signal system designed to represent referents made of dots and dashes would appear to be highly combinatorial. However, such combinatoriality would be spurious, because it would not reflect any organizational principle of the communication system itself (aside from that of representing its referents iconically). This point is clearly relevant to our study — in particular the Mimable-line condition — but it is important to reiterate that full iconicity was not mandated even in the mimable conditions. The key result is that, although combinatorial systems are simpler to produce, participants rarely took advantage of this fact (such as by producing more combinatorial, semi-iconic signals) when full iconicity was available as a strategy. This study thus offers an important complement to that of Roberts and Galantucci (2012). While that study investigated the relationship between combinatoriality and iconicity in conditions where the latter was present but relatively weak, our study investigated the consequences for combinatoriality when the opportunity for iconicity was maximized (and, in the Non-mimable condition, minimized). Finally, it should be noted that while the line referents were not themselves obviously combinatorial, nor were the color referents. The extent to which participants produced combinatorial signals in the Non-mimable condition is thus as striking as the extent to which they adopted fully iconic ones in the Mimable conditions.

So far this paper has concerned itself only with the early stages of communication systems, at which point — as we have shown — the pressure for referential efficiency trumps the pressure for transmission efficiency. However, these two pressures act on any communication system, regardless of its stage of development, and it may well be that at different stages they interact in different ways. It should not be assumed, in other words, that the referential efficiency offered by iconicity will always trump the transmission efficiency offered by combinatoriality. Once signs are well established, the benefits of iconicity may become less compelling, and over time we might expect combinatoriality to emerge at the expense of iconicity. Our game ended too soon for such a process to become apparent. Sandler et al. (2011), however, found evidence in ABSL for the gradual conventionalization of iconic signs and the beginnings of combinatorial structure.

Over time there may also be an increase in the number of meanings that need to be expressed, and this has long been hypothesized to be a driving force behind the pressure for combinatoriality (Hockett, 1960a, 1960b; Lindblom, MacNeilage, & Studdert-Kennedy, 1984; Nowak et al., 1999; Pinker & Jackendoff, 2005). However, like other Experimental-semiotic studies (Del Giudice, 2012; Galantucci et al., 2010; Roberts & Galantucci, 2012) our study did not find support for this hypothesis. Indeed, sign-sets were smaller in the Non-mimable referent condition, in which Combinatoriality was greater, than in either of the other two conditions (Fig. 8a). It is possible, as Galantucci et al. (2010) speculated about their own study, that the number of meanings expressed is below the threshold at which set- size exercises an effect. However, the fact that Combinatoriality emerged in our study when there were only ten meanings to express, and in the condition with the smallest sign-sets, suggests this may simply be the wrong place to look for an explanation for the emergence of combinatoriality.

There may be more than one right place to look. Galantucci et al. (2010) focused on rapidity of fading and found that combinatoriality was greater when signals faded faster; we focused on iconicity and found greater combinatoriality when there was less opportunity for iconicity. Iconicity and rapidity of fading may play independent roles, perhaps alongside other factors. Given that Galantucci et al.'s fast-fading signals were also less iconic, however, it may be that the effect of rapidity of fading on combinatoriality is mediated by iconicity. Moreover, while in our study we manipulated iconicity while keeping modality constant, it remains possible that differences between the vocal-auditory and manual-visual modalities play a role independent from iconicity in explaining variation in combinatoriality. The current study does not allow us to answer these questions, but the Experimental-semiotic approach provides the tools to do so and expand our increasingly nuanced view of the basic design principles of human communication.

Author contributions

BG developed the study concept. All authors contributed to the study design. Testing and data collection were performed by GR and JL. GR performed the data analysis. GR and BG drafted the paper, and JL provided critical revisions.

Acknowledgments

We thank Carrie Theisen and Dario Gutierrez for helpful comments, and Christian Kroos for writing the communication game software. This research was supported by the National Science Foundation (BCS-1026943).

References

Abler, W. L. (1989). On the particulate principle of self-diversifying systems. *Journal of Social and Biological Structures, 12*(1), 1–13.

Burling, R. (1999). Motivation, conventionalization, and arbitrariness in the origin of language. In B. J. King (Ed.), *The origins of language: What nonhuman primates can tell us* (pp. 307–350). School of American Research Press.

Del Giudice, A. (2012). The emergence of duality of patterning through iterated learning: Precursors to phonology in a visual lexicon. *Language and Cognition, 4*(4), 381–418.

Donald, M. (1991). *Origins of the modern mind*. Cambridge, MA: Harvard University Press.

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

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 Psychological Science, 5*, 354.

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

Galantucci, B. (2009). Experimental Semiotics: A new approach for studying communication as a form of joint action. *Topics in Cognitive Science, 1*(2), 393–410.

Galantucci, B., Garrod, S., & Roberts, G. (2012). Experimental Semiotics. *Language and Linguistics Compass, 6*(8), 447–493.

Galantucci, B., Kroos, C., & Rhodes, T. (2010). The effects of rapidity of fading on communication systems. *Interaction Studies, 11*(10), 100–111.

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

Goldin-Meadow, S., & McNeill, D. (1999). The role of gesture and mimetic representation in making language the province of speech. In M. C. Corballis & S. E. G. Lea (Eds.), *The descent of mind: Psychological perspectives on hominid evolution* (pp. 155–172). Oxford: Oxford University Press.

Goldin-Meadow, S., Mylander, C., & Butcher, C. (1995). The resilience of combinatorial structure at the word level: Morphology in self-styled gesture systems. *Cognition, 56*, 195–262.

Hockett, C. F. (1960a). The origin of speech. *Scientific American, 203*, 88–96.

Hockett, C. F. (1960b). Logical considerations in the study of animal communication. In W. Lanyon & W. Tavolga (Eds.), *Animal sounds and communication* (pp. 392–430). Washington: American Institute of Biological Sciences.

Krifka, M. (2001). Compositionality. In R. A. Wilson & F. Keil (Eds.), *The MIT encyclopaedia of the cognitive sciences*. Cambridge, MA: MIT Press.

Ladd, D. R. (2012). What is duality of patterning, anyway? *Language & Cognition, 4*(4), 261–273.

Lindblom, B., MacNeilage, P. F., & Studdert-Kennedy, M. (1984). Self-organizing processes and the explanation of phonological universals. In B. Butterworth, B. Comrie, & O. Dahl (Eds.), *Explanations for language universals* (pp. 181–203). Berlin: Mouton.

Martinet, A. (1960). *Elements of general linguistics*. Chicago: University of Chicago Press.

Meier, R. (2002). Why different, why the same? Explaining effects and non-effects of modality upon linguistic structure in sign and speech. In R. Meier, K. Cormier, & D. Quinto-Pozos (Eds.), *Modality and structure in signed and spoken languages* (pp. 1–25). Cambridge: Cambridge University Press.

Meir, I., Israel, A., Sandler, W., Padden, C., & Aronoff, M. (2012). The influence of community on language structure. *Linguistic Variation, 12*(2), 247–291.

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*, 20130299.

Nowak, M. A., Krakauer, D. C., & Dress, A. (1999). An error limit for the evolution of language. *Proceedings of the Royal Society B — Biological Sciences, 266*(1433), 2131–2136.

Page, D. L., Koschan, A., Sukumar, S., Abidi, B., & Abidi, M. (2003). Shape analysis algorithm based on information theory. *Proceedings of the International Conference on Image Processing, I*, 229–232.

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*, 1–15.

Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition, 95*(2), 201–236.

Roberts, G., & Galantucci, B. (2012). The emergence of duality of patterning: Insights from the laboratory. *Language and Cognition, 4*(4), 297–318.

Sandler, W., Aronoff, M., Meir, I., & Padden, C. (2011). The gradual emergence of phonological form in a new language. *Natural Language & Linguistic Theory, 29*(2), 503–543.

Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences of the United States of America, 102*(7), 2661–2665.

Senghas, A., Özyürek, A., & Goldin-Meadow, S. (2010). The evolution of segmentation and sequencing: Evidence from Homesign and Nicaraguan Sign Language. In A. D. M. Smith, M. Schouwstra, B. de Boer, & K. Smith (Eds.), *The evolution of language: Proceedings of the 8th International Conference (EVOLANG8)* (pp. 279–288). Singapore: World Scientific.

Stokoe, W. C. (1991). Semantic phonology. *Sign Language Studies, 70*, 107–114.

Studdert-Kennedy, M. (2000). Evolutionary implications of the particulate principle: Imitation and the dissociation of phonetic form from semantic function. In C. Knight, M. Studdert-Kennedy, & J. R. Hurford (Eds.), *The evolutionary emergence of language* (pp. 161–176). Cambridge: Cambridge University Press.

Taub, S. F. (2001). *Language from the body*. Cambridge: Cambridge University Press.

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

Tria, F., Galantucci, B., & Loreto, V. (2012). Naming a structured World: A cultural route to duality of patterning. *PLoS ONE, 7*(6), e37744.

Verhoef, T. (2012). The origins of duality of patterning in artificial whistled languages. *Language and Cognition, 4*(4), 357–380.

Verhoef, T., Kirby, S., & de Boer, B. (2013). Combinatorial structure and iconicity in artificial whistled languages. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th annual conference of the cognitive science society* (pp. 3669–3674). Berlin: Cognitive Science Society.

Wray, A., & Grace, G. W. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form". *Lingua, 117*, 543–578.