Contextual predictability shapes signal autonomy

Citation for published version:
Winters, J, Kirby, S & Smith, K 2018, 'Contextual predictability shapes signal autonomy' Cognition, vol. 176, pp. 15-30. DOI: 10.1016/j.cognition.2018.03.002

Digital Object Identifier (DOI):
10.1016/j.cognition.2018.03.002

Published In:
Cognition

Contextual predictability shapes signal autonomy

James Winters[1,1,*], Simon Kirby, Kenny Smith

Abstract

Aligning on a shared system of communication requires senders and receivers reach a balance between *simplicity*, where there is a pressure for compressed representations, and *informativeness*, where there is a pressure to be communicatively functional. We investigate the extent to which these two pressures are governed by *contextual predictability*: the amount of contextual information that a sender can estimate, and therefore exploit, in conveying their intended meaning. In particular, we test the claim that contextual predictability is causally related to *signal autonomy*: the degree to which a signal can be interpreted in isolation, without recourse to contextual information. Using an asymmetric communication game, where senders and receivers are assigned fixed roles, we manipulate two aspects of the referential context: (i) whether or not a sender shares access to the immediate contextual information used by the receiver in interpreting their utterance; (ii) the extent to which the relevant solution in the immediate referential context is generalisable to the aggregate set of contexts. Our results demonstrate that contextual predictability shapes the degree of signal autonomy: when the context is highly predictable (i.e., the sender has access to the context in which their utterances will be interpreted, and the semantic dimension which discriminates between meanings in context is consistent across communicative episodes), languages develop which rely heavily on the context to reduce uncertainty about the intended meaning. When the context is less predictable, senders favour systems composed of autonomous signals, where all potentially relevant semantic dimensions are explicitly encoded. Taken together, these results suggest that our pragmatic faculty, and how it integrates information from the context in reducing uncertainty, plays a central role in shaping

*Corresponding author
Email address: `winters@shh.mpg.de` (James Winters)

language structure.

Keywords: Language evolution, interaction, communication games, context, pragmatics

---

1. Introduction

Reducing uncertainty about the intended meaning is fundamental to any good communication system (Piantadosi et al., 2012; Ramscar & Port, 2015). In achieving this aim, speakers and hearers need to coordinate with one another, relying not only on the creation of conventional forms, but also on the way these forms interact with the contextual information at hand (Lewis, 1969; Sperber & Wilson, 1986; Clark, 1996; Croft, 2000; Scott-Phillips, 2015). Without context, linguistic systems such as English would be woefully ambiguous, leaving the sentence *She passed the mole* uninterpretable as to whether the verb *passed* refers to *a form of motion* or *an act of giving* and whether the noun *mole* refers to *a small burrowing mammal*, *a person engaged in espionage*, *a Mexican sauce* or *a type of causeway*. In short, when the context is known and informative, it helps in reducing uncertainty (Piantadosi et al., 2012).

Context, in this sense, refers to the mutual cognitive environment (Sperber & Wilson, 1986) in which an utterance is situated and determines what is and is not informative for reducing uncertainty. This consists of a figure (the target of interpretation), a ground (the immediate information brought to the act of interpretation), and a background (prior knowledge derived from previous frames) (Duranti & Goodwin, 1992; Terkourafi, 2009). And, as with any environment, the context will vary: some contexts are regular and predictable, whereas others fluctuate and are unpredictable. When viewed in this way, the context is a variable that determines the extent to which a speaker can estimate, and therefore exploit, information that is relevant to reducing uncertainty about their intended meaning – its *contextual predictability.*

For instance, if a speaker is providing directions to the nearest grocery store, then the context includes information in the immediate environment, such as the general direction of the store relative to the present position of the interlocutors, as well as background knowledge about how a hearer is likely to interpret an utterance given the outcomes of previous interactions. Predictable contexts are therefore those where the speaker is able to use information provided by the context to reduce uncertainty about their intended meaning: if the grocery store is near a park, and the speaker and hearer share knowledge about where this park is located, then saying "there's a grocery store about five minutes away, next to the park where we play rugby" is sufficient for the hearer to find the grocery store. This is in contrast to a situation where the speaker and hearer are strangers and uncertainty exists as to the knowledge they both share with one another (e.g., the hearer is a tourist and does not know about the existence of a nearby park).

This relationship between context, meaning and uncertainty leads to an interesting trade-off in how linguistic systems are organised. Languages vary in their degree of *signal autonomy*: "the capacity for an utterance to be interpreted in isolation, without recourse to implicit linguistic, cultural, contextual or cotextual knowledge. Non-autonomous expression combines linguistic signals with context, pragmatics, paralinguistic signals and the like" (Wray & Grace, 2007: 556). One hypothesis is that autonomy is favoured in situations where speakers and hearers cannot rely on context for disambiguation (Kay, 1977): autonomous signals are advantageous insomuch as they reduce reliance on shared social, historical and local contexts in favour of internal structure (Snow et al., 1991; Hurford, 2011).

In this paper we present experimental evidence demonstrating that the degree of signal autonomy is causally related to contextual predictability: in an experiment where participants interact using an artificial language, highly predictable contexts favour systems composed of non-autonomous, context-dependent signals, whereas decreasing contextual predictability results in increased autonomy (context-independence). Crucially, we argue that these systems arise from the pressures of *informativeness* and *simplicity* (Regier et al., 2015; Kirby et al., 2015), with the degree of contextual predictability interacting with these two pressures to restrict the space languages explore.

1.1. Signal Autonomy and Contextual Predictability

No natural language has completely autonomous signals in the sense of unambiguous clarity; context is always involved in reducing uncertainty about the intended meaning. But it is relatively uncontroversial to say there are degrees of autonomy. Contrast the use of indexical (context-dependent) and non-indexical (autonomous) forms of language: when referring to the day after today, English users will tend to say *tomorrow*, rather than the more autonomous counterpart of a specific date (e.g., *July 5th 2016*) (Hurford, 2011). Both are perfectly valid forms of expressing the relevant meaning, yet indexical forms require enrichment from external information (e.g., James lives on *this* street), whereas non-indexical forms are useful in providing specific information in the absence of such contexts (e.g., James lives on Milton Street).

It is not just individual constructions which vary in signal autonomy; languages, as collections of constructions, also vary in the extent to which they can be characterised as more or less autonomous ((Kay, 1977); Wray & Grace, 2007; Hurford, 2011). An extreme example of this cross-linguistic variation in autonomy is found in Riau Indonesian — a colloquial variety of Malay/Indonesian with minimal syntactic structure and highly context-dependent expressions (for review see Gil, 2005). For instance, the combination of *ayam* ("chicken") and *makan* ("eat") yields a vast number of possible interpretations, ranging from *the chicken is eating* to *the chickens are eating* or *someone is eating the chicken* or even *someone is eating with the chicken* (Gil, 2005; Hurford, 2011). In short, the phrase *ayam makan* or *makan ayam* involves anything to do with chicken and eating; contextual information and inference do the rest of the work in sifting through possible interpretations.

1.2. The immediate context, the amount of shared context, and the historical context

Our general hypothesis is that variation in signal autonomy is predicted by the amount of contextual predictability. When the context is predictable, signal autonomy is low; when the context is unpredictable, signal autonomy is high. A complicating factor is determining what aspects of the context shape the degree of contextual predictability. This is problematic because operationalising separate aspects of context, and investigating the relationships between these aspects in a systematic fashion, is no simple task (see Clark & Carlson, 1981; Bazire & Brézillon, 2005). For our purposes, we restrict our focus to three types of context: *the immediate context*, *the amount of shared context*, and *the historical context*.

The immediate context refers to the situational information that is relevant for producing and comprehending an utterance. Consider the possible use of referring expressions in Figure 1. Describing the object on the left in contexts A and B could be achieved with the referring expression *the metal cup* – this expression is capable of discriminating between referents in both contexts. Yet, based on a long history of psycholinguistic studies, it is only in context B where the expression *the metal cup* is used, with *the cup* being preferred when the adjective is not needed for discrimination (Olson, 1970; Pechmann, 1989; Sedivy, 2005; for review, see: Konopka & Brown-Schmidt, 2014). In this sense, a maximally predictable context is one in which a single semantic dimension (e.g., shape) is relevant for discrimination, with predictability decreasing as more dimensions are necessary for success (e.g., shape and material).

Figure 1: In context A, an English speaker can discriminate between both objects by using *the cup* or *the bowl*, whereas in context B they must use more elaborate expressions: *the metal cup* and *the wooden cup* (assuming the speaker obeys the rules of English for adjective use).

If the immediate context is the only relevant factor in determining contextual predictability, signal autonomy should pattern with the amount of contextual information in the local, situational context. In this case, low signal autonomy is expected when the situational context backgrounds some information (e.g., material) and highlights other information (e.g., shape). This narrow conception of context runs into problems when accounting for linguistic phenomena such as overspecification (where redundant, non-contrastive information is incorporated; see: Tinits et al., 2017). For instance, unlike material and scalar adjectives, which tend to pattern with the immediate context, colour adjectives are often used even when they are uninformative for discrimination[1] (e.g.,

[1]A growing body of work into these *Redundant Colour Adjectives* (RCA) provides two explanations (Rubio-Fernández, 2016). First, the use of RCAs tends to be contingent on the semantic category, as evident in their presence for atypical objects (e.g., *the brown banana*) and clothes (e.g., collocations such as *black tie*) and their absence in typical (e.g., *the banana*) and geometrical figures (Dale & Reiter, 1995; Grodner & Sedivy, 2011). Second, speakers tend to produce RCAs when colour helps facilitate object recognition (e.g., polychrome versus monochrome displays), as well as when the language uses pre-nominal (e.g., English) as

Sedivy, 2005; Arts et al., 2011; Rubio-Fernández, 2016). Similarly, languages often morphologically overspecify, obligatorily encoding distinctions which are unnecessary to the current discourse topic[2] (Lupyan & Dale, 2010; Trudgill, 2011).

In addition to the immediate context, the generation and interpretation of utterances are also constrained by the amount of knowledge shared between a speaker and a hearer (e.g., mutual cognitive environment: Sperber & Wilson, 1986; common ground: Clark & Carlson, 1981). Speakers often produce longer, more elaborate expressions when the hearer is perceived to be less knowledgeable about a topic (Isaacs & Clark, 1987; Heller et al., 2009), and the use of scalar-modified expressions is partially contingent on whether or not a speaker and a hearer share the same referential context (Nadig & Sedivy, 2002). The extent to which speakers actively take into account the informational needs of their audience is strongly debated. Clark (1992: 80-81), for instance, argues that speakers adhere to a *Principle of Optimal Design*, whereby the speaker designs his or her utterances on the basis of the expectations and beliefs of the hearer. Work by Keysar and colleagues (Horton & Keysar, 1996; Keysar et al., 1998; Horton & Gerrig, 2005) demonstrates that this is not always the case: speakers often start out egocentric, in that they primarily consider their own perspective when designing utterances, and only consider informational needs of their partner after this initial stage of processing (e.g., when communication breaks down).

Still, a basic assumption shared by all these studies is that the amount of information shared by interlocutors impacts communication, with more shared contextual information reducing uncertainty about the speaker's intended meaning. In this respect, less shared contextual information is associated with a

---

opposed to post-nominal adjectives (e.g., Spanish) (Rubio-Fernández, 2016).

[2]For example, Yagua, a language of Peru, differentiates 5 levels of remoteness through inflectional synthesis of the verb (see Lupyan & Dale, 2010). English can express remoteness distinctions lexically (contrast "I went to the shops yesterday" with "I went to the shops two weeks ago"), but expressing this information is optional; Yagua speakers, by contrast, must obligatorily encode remoteness as an inflection.

reduction in contextual predictability, as a speaker is unable to estimate and exploit information that is useful for a hearer. A reduction in shared contextual information must be compensated for by an increase in signal autonomy. Figure 2 provides a simple illustration of this point: here, the goal is for the speaker to convey the same intended meaning in two different contexts (A and B). Context A has referents that share the same colour but differ in shape, and context B has referents that share the same shape but differ in colour. In the shared context, the speaker and hearer both have access to the same contextual information, whereas in the unshared context the speaker only sees the referent they need to convey. Assuming speakers have knowledge about a set of signals, and are free to combine them, then there are three possibilities for signalling the target referent: *the blue one*, *the square*, *the blue square*. If signal autonomy is determined by the amount of shared context, then speakers should use *the square* in context A and *the blue one* in context B. However, in the unshared context, speakers only have access to the target they need to convey, and are therefore unable to condition their signals on contextual information. By using the more autonomous form, *the blue square*, speakers can be sure to convey their intended meaning across both contexts.

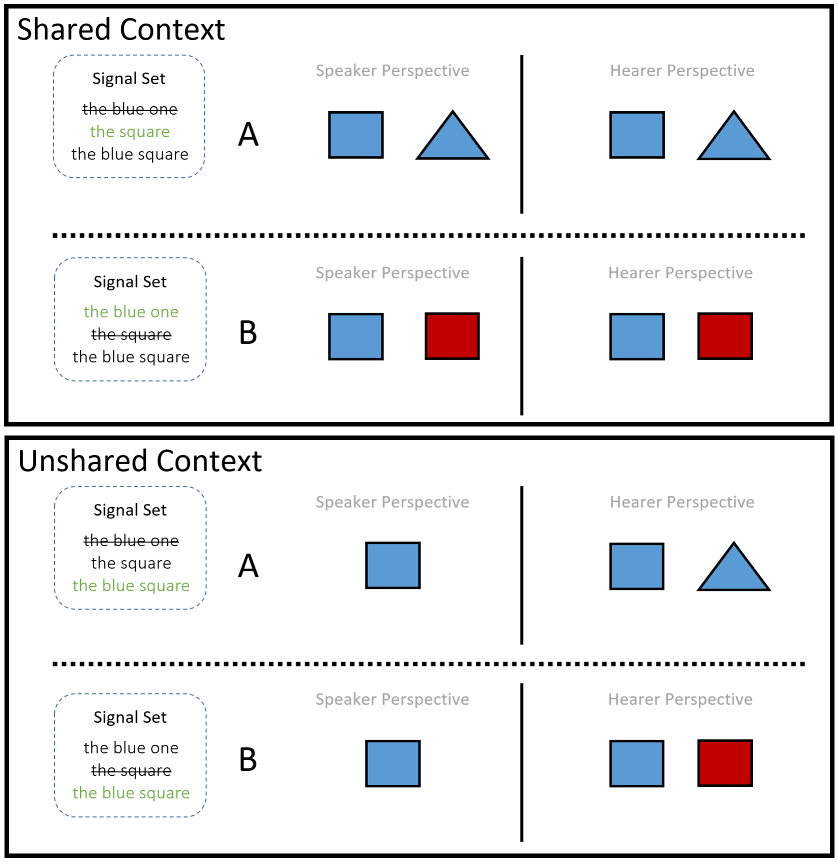


Figure 2: An example of a reference language game where the speaker has a set of signals for conveying a particular meaning (in this case, a blue square). There are two contexts: A and B. In context A, the target image shares the same colour as its distractor (but differs in shape), whereas in context B the target image shares the same shape as its distractor (but differs in shape). For the Shared Context (top box), the speaker has access to the same contextual information as the hearer. By contrast, in the Unshared Context (bottom box), the speaker does not have access to the same contextual information as the hearer. The green coloured signals are the preferred signals for communicating an intended meaning in a particular scenario. The signals with a strikethrough are those which will not convey the intended meaning in a particular scenario.

What enters into a context is also contingent on information extracted from prior events, i.e., the *historical discourse context* (Brennan & Clark, 1996; Van Der Wege, 2009; Yoon et al., 2016). A key insight is that speakers continue to use pre-established, conventional referring expressions even when there is the option to use a more or less autonomous form (Brennan & Clark, 1996). This creates a causal link where the aggregation of previous contexts, and the utterances generated in those contexts, constrain the solutions required in the current context. As such, historical precedent can override novelty when these conventions are generalisable (to new contexts) and functional (in that an expression is capable of identifying the intended meaning).

Figure 3 shows how generalisability is to some extent a function of the historical relationship between contexts (Van Der Wege, 2009; Winters et al., 2015; Yoon et al., 2016). If contexts are predictable, then a solution at an earlier context is generalisable to all subsequent contexts. This is the case for the first sequence: the set of possible contexts always consist of referents from distinct semantic categories. Using *the t-shirt* is sufficient for conveying the target referent across the set of aggregate contexts. A decrease in contextual predictability means that a solution at an earlier context is not necessarily generalisable to another context. This is evident in the second sequence: the set of possible contexts is variable in that the target referent is sometimes paired with an object from a different semantic category and at other times it is paired with an object from the same semantic category. So, even though the speaker is able to refer to the target as *the t-shirt* in the first context, when confronted with the second context they are required to use a more autonomous form (*the striped t-shirt*). In the third context, however, speakers often continue to use the more autonomous, overspecified form, *the striped shirt*, even though *the shirt* is sufficient. This example demonstrates that history can override what is locally relevant: prior solutions become "locked-in", established as conventions, and introduce contingencies into future outcomes (Millikan, 1998).

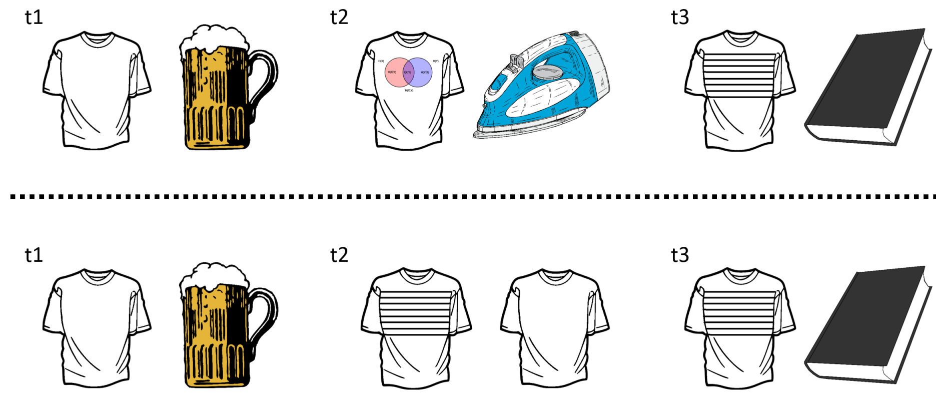


Figure 3: An example of two sequences to show how the historical relationships between referential contexts can influence the outcome of utterance use. The target image is always the object on the left. Sequence one (top) shows target-distactor pairings that always differ in their semantic category. Speakers will overwhelmingly use *t-shirt* to convey the target as it is generalisable from *t1* to *t3* (overspecification in these cases tends to be low; see Yoon et al., 2016). In sequence two (bottom), target-distractor pairings sometimes differ in their semantic category (as in *t1*) and sometimes they share the same semantic category (as in *t2*). This is important because at *t3* we predict that speakers will continue using an overspecified utterance, *the striped t-shirt*, even though *t-shirt* is sufficient for conveying the target (Brennan & Clark, 1996; Van Der Wege, 2009).

1.3. Competing pressures: Simplicity and Informativeness

Contextual predictability, then, can be seen as an organising principle for how we use and structure language: when the context is highly predictable, use less autonomous forms, and in less predictable contexts use more autonomous forms. However, understanding the causal relationship between contextual predictability and language structure requires we take seriously the *problem of linkage* (Kirby, 1999): How do the behaviours of individuals give rise to the particular structural properties of language?

One solution to the problem of linkage is to consider how short-term strategies used in solving immediate communicative needs can give rise to language systems through long-term patterns of learning and use (Evans & Green, 2006; Beuls & Steels, 2013; Winters et al., 2015). Shaping these short-term strategies are competing motivations of speakers and hearers (Zifp, 1949; Frank & Goodman, 2012; Piantadosi et al., 2012; Kemp & Regier, 2012; Kirby et al., 2015). Here, a balance needs to be reached between *simplicity*, where there is a cognitive pressure for compressed representations (Chater & Vitányi, 2003; Culbertson & Kirby, 2016), and *informativeness*, where there is a communicative pressure for accuracy (Piantadosi et al., 2012; Kirby et al., 2015). Resolving these two pressures require speakers and hearers to align on a system that reaches a tradeoff between compressible signal-meaning mappings whilst also reducing uncertainty about the intended meaning. Simpler semantic representations are tolerated inasmuch as our pragmatic capacity can make use of contextual information to fill in expressive gaps in a language (Brochhagen et al., 2016).

The pressures imposed during communication provide a clear prediction linking together signal autonomy and context. Non-autonomous signals are expected when contextual predictability is high. Simple, context-dependent signal-meaning mappings can encode less information because contextual information is predictable enough to enrich the speaker's intended meaning. These context-dependent systems are simpler because they make use of fewer signals for achieving successful communication. By contrast, a decrease in contextual predictability is expected to result in more autonomous signals. As speakers cannot rely on contextual information to convey their intended meaning, encoding more information provides a robust communicative solution: autonomous signals are less sensitive to changes in contextual information, increasing the probability of a hearer arriving at the correct interpretation at the expense of a simpler system (i.e., autonomous systems require more signals).

Several recent studies directly link the effects of context to the emergence of structure (Silvey et al., 2014; Winters et al., 2015; Tinits et al., 2017). Communication games are a useful tool for exploring the dynamics underpinning communicative behaviour in a laboratory setting (for review, see: Galantucci et al., 2012; Tamariz, 2017). In contrast to standard referential games, where participants already have a rich, highly structured and conventional language for communication, these studies use artificial languages to study the *de novo* emergence of communication systems. This allows experimenters to control away certain features (such as knowledge of a shared vocabulary) and force participants to negotiate a shared system from scratch. Two general findings emerge from this literature with regards to contextual predictability and signal autonomy: first, languages do optimise when the context always highlights one relevant semantic dimension (e.g., shape) and backgrounds the irrelevant semantic dimension(s) (e.g., colour), with the system adapting to only encode the dimension relevant to discriminating between referents (Silvey et al., 2014; Winters et al., 2015); second, languages are more autonomous when the relevant semantic dimension varies across the set of contexts (e.g., shape is relevant in one context, yet colour is relevant in another).

We extend these findings by manipulating both the access to a shared context and the historical context to test for the effect of contextual predictability in determining signal autonomy. This allows us to overtly investigate the relationship between short-term strategies in the immediate referential context and the emergence of language structure across the set of aggregate contexts.

1.4. Manipulating Contextual Predictability

In our experiment, participants are first trained on an initially ambiguous artificial language, and then placed in an *asymmetric communication game* (Moreno & Baggio, 2015) where they are assigned fixed roles as either sender or receiver. Sender and receiver play a series of guessing games (Steels, 2003): the task is for the receiver to discriminate between a target object and a set of distractor objects using a signal provided by the sender. Possible referents were drawn from a set of images which vary in shape and colour. The training language is ambiguous in that it underspecifies on whether labels encode shape, colour or both (see below for details of how this is achieved). This allows us to explore how senders convey novel meanings, and how much information they choose to encode explicitly in the linguistic signal.

To test for the effect of contextual predictability on autonomy, we made two manipulations: (i) Access to context (Shared Context/Unshared Context), and (ii) Context-type (Shape-Different/Mixed). Access to context varies the amount of contextual information shared between sender and receiver. In the *Shared Context* conditions, senders have access to the context against which their utterance will be interpreted (i.e., the array of target and distractors that the receiver is confronted with), whereas in the *Unshared Context* conditions senders only see the target in isolation (although the sender's task remains the same: to produce a signal which allows the receiver to distinguish a target from a set of distractors). Senders in Shared Context conditions therefore have knowledge about what distinctions they need to make on a trial-by-trial basis, whereas senders in Unshared Context conditions only know what target they need to convey (without any contextual information about the context against which their utterance will be interpreted, and therefore what the relevant distinctions are for the receiver in a particular trial).

Our second manipulation of contextual predictability involves context type: to what extent is a particular semantic dimension (e.g., shape) relevant for discrimination across successive trials? This allows us to test for the effect of the historical context in determining signal autonomy. For the *Shape-Different* conditions, the context-type remains consistent across trials, with targets and distractors always differing in shape (but sharing the same colour). *Mixed* conditions vary their context-type across trials: half of the trials consist of contexts in which the target and distractors differ in shape (but share the same colour) and half in which they differ in colour (but share the same shape). In Shape-Different conditions, encoding shape is therefore always sufficient to allow the receiver to retrieve the intended meanings; in Mixed conditions, some trials will require the encoding of shape, some will require colour to be encoded. This gives us four conditions: *Shape-Different + Shared Context*, *Shape-Different + Unshared Context*, *Mixed + Shared Context*, *Mixed + Unshared Context*.

In terms of contextual predictability, the Shape-Different + Shared Context condition is the most predictable both within and across trials: the context-type is consistent, in that Shape is always the relevant feature for discrimination, and the sender has access to the same contextual information as the receiver. The simplest solution here is for senders to use the contextual information to only encode shape in their signals, resulting in a system with low signal autonomy (out of context a signal has a decreased capacity to discriminate between referents). On the opposite end is the Mixed + Unshared Context condition: context-type varies between trials, with access to this contextual information being unavailable for the sender (they only ever see the target that needs to be conveyed). This low contextual predictability means that underspecified systems will be ineffective — in order to be sure of conveying their intended meaning, senders must instead employ strategies that increase signal autonomy, e.g., by encoding both shape and colour on every trial.

For the Shape-Different + Unshared Context and Mixed + Shared Context conditions there is one manipulation which decreases contextual predictability and another which increases contextual predictability. We therefore expect heterogeneous outcomes in these two conditions. In the Shape-Different + Unshared Context condition, the fact that the sender lacks access to the context favours strategies that increase signal autonomy, as the sender has no contextual information regarding what distinctions they need to convey. However, the across-trial predictability potentially allows senders to reduce their signal autonomy, as encoding shape is always sufficient for conveying the intended meaning. A similar story holds for the Mixed + Shared Context condition. This time the variability across trials decreases contextual predictability; however, the fact that the sender has access to the same contextual information as the receiver should increase contextual predictability. Whether or not a sender opts for strategies that increase or decrease signal autonomy is contingent on the initial assumptions a sender brings to the task, the feedback they receive from receivers and the relative contributions of within-trial and across-trial contextual predictability.

2. Method

2.1. Participants

120 undergraduate and graduate students at the University of Edinburgh (79 female, 41 male, median age 20) were recruited via the Student And Graduate Employment database and randomly assigned to one of the four possible conditions (see § 2.3.3). Each condition consisted of a pair of participants who learned an artificial language (see § 2.2) and then used this language in a communication game (see § 2.3.2). Participants were paid £5 for their participation.

2.2. Stimuli: Images and Target Language

Participants were asked to learn and then use an 'alien language', consisting of lower-case labels paired with images. There were 16 images that varied along three features: shape, colour and a unique identifier (see figure 4 for examples). Four of these 16 images were randomly selected for training, such that each colour and shape was represented exactly once and each of the four images therefore differed from all the others in both colour and shape. Each image was then assigned a label as follows: From a set of vowels (a,e,i,o,u) and consonants (g,h,k,l,m,n,p,w) we randomly generated nine CV syllables which were then used to randomly generate a set of four 2-3 syllable words (e.g., *kewa*, *nunuki*, *lono*, *mopola*). Since the four images used during training differed in both shape and colour, the training labels in this language were therefore ambiguous with respect to whether they referred to colour, to shape, or to both colour and shape (or, equivalently, the unique identifier).

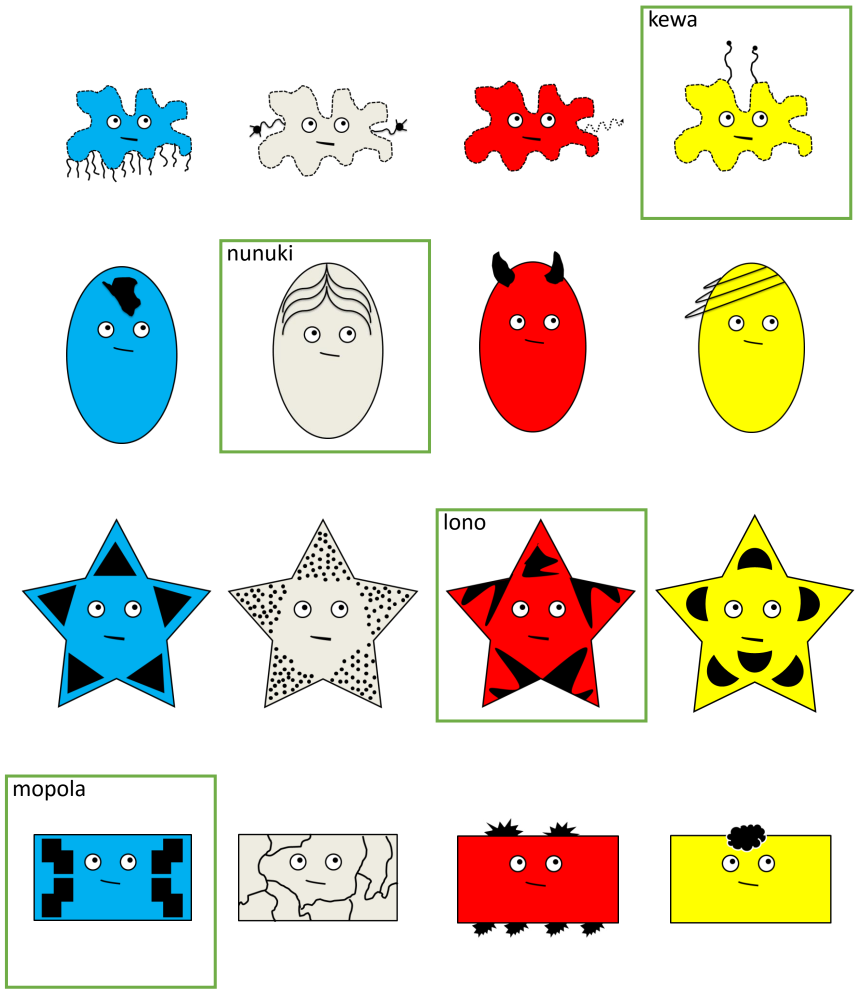


Figure 4: The referent space used in our experiment. Referents vary along three dimensions: a shape, a colour, and a unique identifier. To create an initially ambiguous language we first randomly selected four referents that differed from one another on both shape and colour. We then paired these four referents with randomly generated signals (see signals and referents inside green boxes).

A total of 15 training languages were created for and shared across the four conditions, with one participant in each condition being exposed to each training language. Yoking participants in this way allows us to control for the effect of initial language on task performance.

2.3. Procedure

Participants were scheduled to participate in pairs. At the start of the experiment, participant pairs were told they would first have to learn and then communicate using an alien language. Participants completed the experiment in separate booths on networked computers. The experiment consisted of two main phases: a *learning phase* and a *communication phase*. Before each phase began, participants were given detailed information on what that phase would involve and were explicitly told not to use English or any other language they knew during the experiment[3]. For the learning phase, participants were trained separately, and it was only during the communication phase that they interacted (remotely, over the computer network).

2.3.1. Learning Phase: Training and Testing

The learning phase was broken up into two components: a set of *training blocks* and a set of *testing blocks*. In each training trial, the participant was presented with an image and a label. After two seconds, the label would disappear, and the participant was then prompted to retype the label before proceeding to the next trial. Each training block consisted of twelve trials (each of the four training images was seen three times, with the order of trials randomised within the block). For the testing trials, the participant was presented with an image and prompted to type the label that corresponded to the image. Once they typed the label, the participant was given feedback as to whether or not they were correct — if incorrect, they were shown the correct label before moving onto the next trial. A testing block consisted of sixteen trials (meaning that each of the four training images was seen four times, with the order of trials randomised within the block). The training and testing blocks were interleaved as follows: 2 training blocks, 1 testing block, 2 training blocks, 1 testing block.

[3]Compliance with the instructions was excellent — there were no cases in which participants used English

2.3.2. Communication Phase

During the communication phase of the experiment, participants within a pair were randomly assigned fixed roles of either *sender* or *receiver* (these roles remained fixed for the during of the communication phase, and did not alternate):

Sender On each communication trial, the sender was shown a target image that was highlighted with a green border. Whether or not a sender could view the distractors depended on the experimental condition (see § 2.3.3). The sender was then prompted to type a description that would best communicate the target to the matcher. Senders were free to type any description they wished (subject to the requirement to avoid English or any other natural language). This description was then sent to the receiver's computer.

Receiver Receivers were presented with an array of four images; the description provided by the sender appearing underneath. Of these four images, one was the target image and the other three images were distractors. Distractors were randomly generated within the constraints imposed by the experimental conditions (see § 2.3.3). The receiver's goal was to click on the image they thought corresponded to the description provided.

Participants were tested on all 16 images during the communication phase, requiring the sender to generalise from the signals provided for the four images in the training set, and the receiver to interpret these generalisations. Following each trial, both sender and receiver were given feedback as to whether or not the receiver had correctly identified the target image described by the sender: both participants were simply informed whether the receiver was correct or incorrect. The communication phase was comprised of three blocks, with each block consisting of 32 trials (trial order was randomised and each of the 16 images appeared as the target image twice within a block).

2.3.3. Manipulating Access to Context and Context-Type

During communication, we manipulated two variables associated with the referential context: (i) access to the referential context; (ii) the referential context type.

The first manipulation consisted of whether or not a sender had access to the referential context that the receiver saw. In the *Shared Context* conditions, senders were exposed to the same referential context as the receiver: that is, they had access to an array consisting of the target and its distractors. Conversely, for the *Unshared Context* conditions, senders only had access to the target image. We manipulated context type (our second manipulation) by varying the possible combinations of target and distractor images within a single trial. For the *Mixed* conditions, half of all trials consist of referential contexts in which the target and its distractors have different shapes (but share the same colour) and half in which they have different colours (but share the same shape). For the *Shape-Different* conditions, the referential context-type remains consistent across trials, with the target and distractors having different shapes (but sharing the same colour) (see Fig. 5).

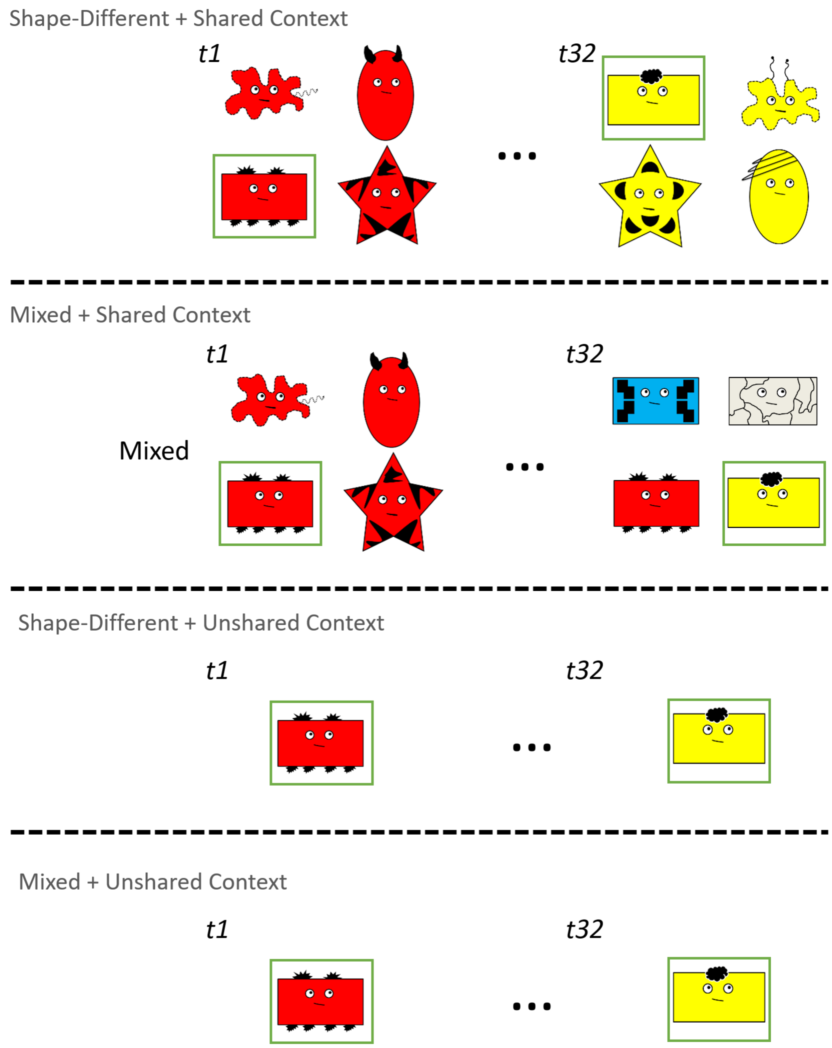


Figure 5: An example of the trial structure for Senders in each of the four conditions. Every block contained 32 trials (i.e., each of the 16 images appeared twice as a target). For Shared Context conditions, the sender always has access to the target (image in green box) and the associated distractors. Conversely, in the Unshared Context conditions, the sender only ever sees the target in isolation. The second manipulation, context-type, is demonstrated in the the top two images. For trials in Shape-Different conditions, the target is always differentiated from the distractors based on shape; For trials in Mixed conditions, the target is sometimes differentiated from the distractors based on shape, whereas on other trials colour is the key differentiating dimension.

2.4. Dependent Variables and Hypotheses

2.4.1. Communicative success

To measure communicative success we recorded the number of successful trials between the sender and receiver, i.e., when the receiver clicked on the target image. The maximum success score was 96 points for three blocks of 32 trials. The purpose of this measure is to see whether the communication systems which develop during interaction are communicatively functional. We therefore predict that all conditions will reach a communicative success score higher than chance (>25%).

2.4.2. Total number of signals

One way to distinguish between autonomous and context-dependent systems is count the total number of unique signals produced. To convey all 16 referents an autonomous system requires more unique signals (16) than a context-dependent system (where the same signal can be reused to express different referents). We predict that the number of signals will be closer to 4 when context is maximally predictable (as in Shape-Different + Shared Context) and closer to 16 when maximally unpredictable (as in Mixed + Unshared Context).

2.4.3. Measuring uncertainty: conditional entropy

To quantify the types of mappings between signals and referents we measure the *conditional entropy* (Cover & Thomas, 1991) of referents given signals for the speaker's productions during interaction (Winters et al., 2015). This gives us a measure of predictability that can be applied to referent uncertainty. $H(R|S)$ is the expected entropy (i.e., uncertainty) over referents given a signal, and therefore captures referent uncertainty,

$$H(R|S) = -\sum_{s \in S} P(s) \sum_{r \in R} P(r|s) \; log P(r|s) \tag{1}$$

where the rightmost sum is the entropy over referents given a particular signal $s \in S$. $P(r|s)$ is the probability that referent r is the intended referent given that signal s has been produced, and $P(s)$ is the probability that signal s will be produced (for any referent). High referent uncertainty corresponds to low signal autonomy, i.e., out of context a signal is highly uninformative about the intended referent, with a speaker reusing that same signal to convey several referents. By contrast, an autonomous signal should have zero referent uncertainty, as each signal a speaker uses only conveys one referent.

While this measure captures the extent to which signals are autonomous, it does not distinguish between context-dependent and counter-functional ambiguity. For context-dependent ambiguity, contextual information contributes to reducing uncertainty about the intended referent, whereas with counter-functional ambiguity this is not the case. To differentiate these two possibilities we also include a measure of referent uncertainty in context, $H(R|S,C)$:

$$H(R|S,C) = -\sum_{s,c \in S,C} P(s,c) \sum_{r \in R} P(r|s,c) \; logP(r|s,c) \tag{2}$$

The rightmost sum now takes into account the entropy over referents given a particular signal in context $s, c \in S, C$. A context $c \in C$ is an array of four referents taken from set R and is constructed so that each referent shares one feature in common and differs on the other feature, e.g., *shape-different blue* = {*blue blob, blue oval, blue square, blue star*} and *colour-different star* = {*blue star, grey star, red star, yellow star*}.

Our general prediction is that systems will vary in referent uncertainty according to the amount of contextual predictability: maximally predictable contexts (as in the Shape-Different + Shared Context) are expected to have high referent uncertainty and minimally predictable contexts (as in Mixed + Unshared Context) are expected to have low referent uncertainty. For our measure of referent uncertainty in context, we predict that all systems will gradually decrease in entropy over time. A context-dependent system should therefore have high referent uncertainty but low referent uncertainty in context: this difference between referent uncertainty and referent uncertainty in context indicates that signals are hard to interpret in isolation but contextual information helps in identifying the intended referent, i.e., the communication system is functionally adequate in context. Conversely, for an autonomous system, we expect both the referent uncertainty and the referent uncertainty in context to be low, i.e., signals are informative even out of context.

3. Results

Our analyses involved four separate mixed effect models (*lme4*: Bates et al., 2015) based on the dependent variables of (a) communicative success, (b) number of unique signals, (c) referent uncertainty, and (d) referent uncertainty in context. For communicative success, we used a logistic mixed effect model and entered Context-Type[4], Access to Context, and Trial (1...96 — Trial was coded such that model intercepts give performance at trial 1) as fixed effects with interactions. For the three other dependent variables – number of unique signals, referent uncertainty, and referent uncertainty in context – we used linear mixed effect models. Context-Type (Shape-Different or Mixed), Access to Context (Shared Context or Unshared Context), and Block (1, 2 and 3 — Block was coded as a factor to maximise contrast) were again entered as fixed effects with interactions. Block was used instead of trial because our measures of entropy and number of unique signals require us to aggregate across a set of trials in each block. We included random intercepts for Participant and initial training language[5], and by-participant random slopes Trial or Block, and by-language slopes for Block/Trial, Context-Type and Access to Context (following the *keep it maximal* approach: Barr et al., 2013). P-values for the fixed effects in the linear mixed effect model were obtained using the *lmerTest* package (Kuznetsova et al., 2015).

[4]Context-Type and Access to Context were centered fixed effects, coded such that positive values of β indicate higher communicative success/number of signals/entropy in Shape-Different or Shared conditions

[5]Participant and initial training language are crossed random effects. This is because the 15 training languages are distributed across 60 participant pairs.

3.1. *Communicative success*

All conditions show levels of communicative accuracy substantially higher than chance ($> 25\%$). This is confirmed by a logistic mixed effect model, which has a significant intercept indicating performance above chance (see Table 1). The significant positive coefficient for Trial suggests that receivers tend to increase their communicative success over successive trials over time. However, this increase is counteracted by the negative coefficients for the two-way interaction between Trial and Context-Type, as can be seen in the Fig. 6 where the two Shape-Different conditions show relatively little increase over trials. Access to Context and Context-Type are also significant predictors of communicative success: conditions where the sender has access to the receiver's context (Shared Context) and where the context-type remains stable across trials (Shape-Different) leads to higher communicative success (as highlighted by the positive coefficients for Access to Context and Context-Type). Finally, the Shape-Different + Shared Context condition (with the highest contextual predictability) is clearly something of an outlier, and the model indicates a significant Context-Type x Access to Context interaction which shows that communicative success in this condition is higher than we would expect given the independent contributions of a shared and stable context.

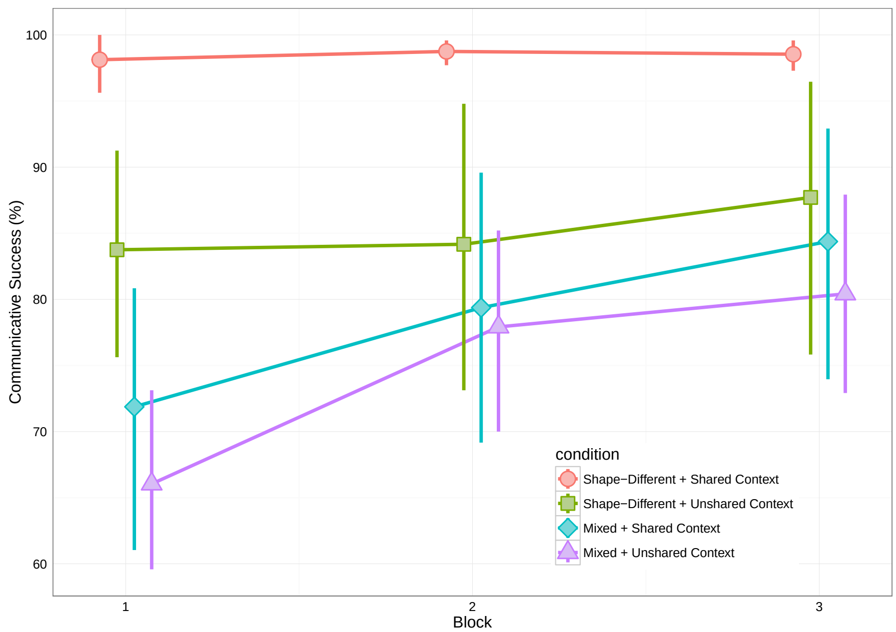


Figure 6: Mean communicative success score by condition over blocks 1-3. Error bars represent bootstrapped 95% confidence intervals. Performance in Shape-Different + Shared Context is near ceiling from block 1, whereas the other three conditions start out at a lower communicative success on average. Both Shape-Different conditions show little improvement in performance. By contrast, the two Mixed conditions tend to show an increase in communicative success from Blocks 1 to 3. All participants scored higher than chance (> 25%) in selecting the target image, indicating that the systems are communicatively functional for identifying the intended referent.

	Estimate β	Std. Error	z value	Pr(>|z|)
(Intercept)	2.182	0.240	33.192	<0.001
Access to Context	1.296	0.426	3.042	0.002
Context-Type	2.684	0.429	6.250	<0.001
Trial	0.010	0.002	4.769	<0.001
Access to Context:Context-Type	2.043	0.851	2.399	0.016
Access to Context:Trial	0.001	0.004	0.171	0.865
Context-Type:Trial	-0.010	0.004	-2.216	0.027
Access to Context:Context-Type:Trial	-0.005	0.009	-0.579	0.563

Table 1: Results for Communicative Success with Access to Context, Context-Type and Trial as fixed effects. The dependent variable is binary (correct or incorrect). The z value and Pr(>|z|) of the Intercept (original z value: 9.077) are adjusted to reflect a chance level of 25%.

3.2. Number of Unique Signals

Figure 7 shows the number of unique signals across condition. As was the case for communicative success, there are significant effects of Context-Type, Access to Context, and a significant interaction for these two predictors: having a shared context and context-types which remain stable over time (i.e., Shape-Different conditions) are associated with smaller signal inventories, suggesting a lesser degree of signal autonomy, and the combination of these manipulations results in very small lexicons in the Shape-Different + Shared Context condition. The marginally significant two-way interaction for Context-Type x Block 3 tells us that the average number of signals decreases from Block 1 to Block 3 in Shape-Different conditions; however, the significant three-way interaction between Access to Context, Context Type and Block 3 suggests that this effect is specific to the Shape Different + Unshared condition (see Table 2).

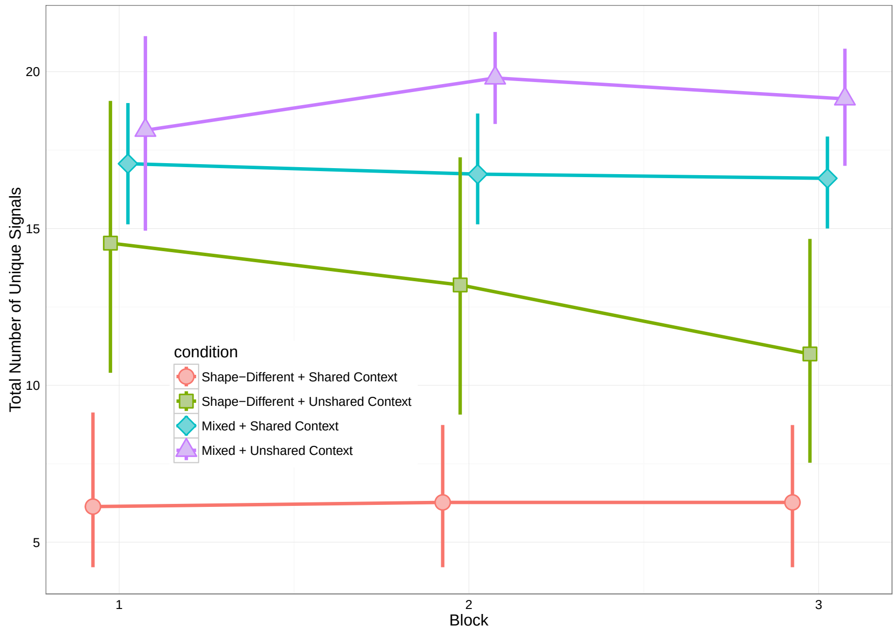


Figure 7: Mean number of unique signals by condition over blocks. Participants in Shape-Different + Shared Context condition use fewer unique signals than participants in the other conditions, and in the Shape-Different + Unshared Context condition the average number of unique signals tends to decrease over time, which suggests the systems are moving away from being autonomous by increasing their context-dependency. By using the minimal number of signals (4) required for successful communication in context, participants in Shape-Different + Shared Context are close-to-optimal in terms of simplicity, whereas participants in Mixed + Shared Context and Mixed + Unshared Context use roughly the number of signals (16) required in an autonomous signalling system. Signalling systems in the Shape-Different + Unshared Context condition appear to be non-autonomous, in that on average participants tend to produce fewer than 16 signals — the number of unique signals changes over the course of interaction. Error bars represent bootstrapped 95% CIs.

	Estimate β	Std. Error	df	t value	Pr(>|t|)
(Intercept)	13.967	0.724	80.454	19.292	<0.001
Access to Context	-4.733	1.448	80.454	-3.269	0.002
Context-Type	-7.267	1.448	80.454	-5.019	<0.001
Block 2	0.033	0.524	112.000	0.064	0.949
Block 3	-0.717	0.524	112.000	-1.367	0.175
Access to Context:Context-Type	-7.333	2.896	80.454	-2.532	0.013
Access to Context:Block 2	-0.267	1.049	112.000	-0.254	0.800
Access to Context:Block 3	1.100	1.049	112.000	1.049	0.297
Context-Type:Block 2	-1.267	1.049	112.000	-1.208	0.230
Context-Type:Block 3	-1.967	1.049	112.000	-1.875	0.063
Access to Context:Context-Type:Block 2	3.467	2.098	112.000	1.653	0.101
Access to Context:Context-Type:Block 3	5.133	2.098	112.000	2.447	0.016

Table 2: Results for Number of Signals with Access to Context, Context-Type and Trial as fixed effects. The dependent variable is continuous.

3.3. Conditional Entropy: referent Uncertainty

Figure 8 plots the conditional entropy of referents given signals, $H(R|S)$, against condition. As a visual inspection of the plot suggests, both Context-Type and Access to Context are significant predictors of referent uncertainty: when contextual predictability increases so too does the out-of-context ambiguity as measured by referent uncertainty, indicating that higher contextual predictability leads to lower signal autonomy. Again, the significant interaction between these two predictors indicates that the combination of a shared and stable context in the Shape-Different + Shared condition produces systems of even lower autonomy than we would expect through the independent contributions of either factor alone. Finally, there is a significant interaction between Context-Type and Block, which runs counter to the significant negative coefficients of Blocks 2 and 3 (see Table 3). This shows that Context-Type influences the evolution of the signalling systems over successive blocks: while referent uncertainty shows a steady decrease in Mixed conditions, it is flat or even increasing in the Shape-different conditions.

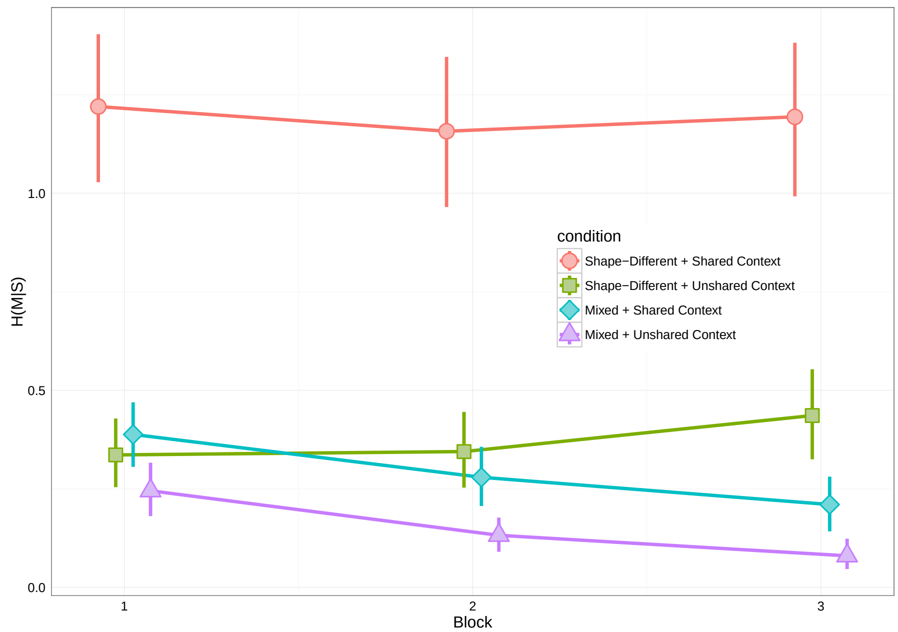


Figure 8: Mean referent uncertainty, measured as $H(R|S)$, by condition over blocks 1-3. Higher entropy values indicate greater out-of-context ambiguity, i.e., lower signal autonomy. Whereas participants in Mixed conditions tend to decrease their referent uncertainty over time, those in Shape-Different conditions tend to either have a high referent uncertainty throughout (as in Shape-Different + Shared Context) or gradually increase their referent uncertainty (as in Shape-Different + Unshared Context). The higher referent uncertainty in Shape-Different + Shared Context suggests participants in this condition are reusing the same signals to express multiple referents. The low referent uncertainty for Mixed + Unshared Context indicates that participants in this condition use each signal to express fewer referents, i.e., senders are producing signals which are unambiguous even out of context.

	Estimate β	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.632	0.078	58.045	8.083	<0.001
Access to Context	0.398	0.155	58.322	2.574	0.013
Context-Type	0.775	0.150	59.302	5.153	<0.001
Block 2	-0.093	0.023	2404.390	-3.967	<0.001
Block 3	-0.115	0.024	2404.771	-4.879	<0.001
Access to Context:Context-Type	0.777	0.298	59.906	2.610	0.011
Access to Context:Block 2	-0.009	0.047	2404.299	-0.191	0.849
Access to Context:Block 3	-0.060	0.048	2404.831	-1.248	0.212
Context-Type:Block 2	0.139	0.049	2404.211	2.852	0.004
Context-Type:Block 3	0.239	0.050	2405.166	4.783	<0.001
Access to Context:Context-Type:Block 2	-0.136	0.101	2404.157	-1.343	0.180
Access to Context:Context-Type:Block 3	-0.164	0.103	2405.211	-1.598	0.110

Table 3: Results for Referent Uncertainty, measured as $H(R|S)$, with Access to Context, Context-Type and Block as fixed effects. The dependent variable is continuous.

3.4. Conditional Entropy: Referent Uncertainty in Context

For the conditional entropy of referents given signals in context, the low entropy values tells us that all of the communication systems are relatively good at identifying the intended referent in context (even at block 1). However, the statistical analysis reveals unexpected differences between conditions: that is, conditions where the sender does have access to the receiver's context (Shared) and the context-type consistently discriminates on the basis of shape (Shape-Different) are, on average, more likely to produce languages with lower levels of uncertainty about the intended referent in context (see Table 4). When contextual predictability is high, this allows systems to reach a simple configuration for communication, with many of systems in Shape-Different + Shared Context reaching zero entropy (i.e., no uncertainty about the intended referent in context). For the other three conditions, where contextual predictability is lower, systems tend to be more complex (see figure 9).

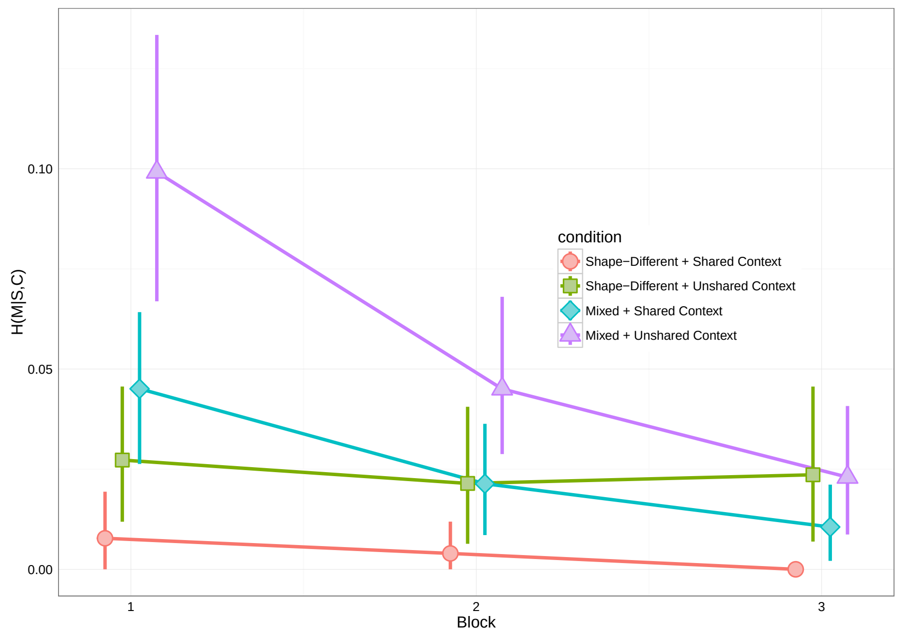


Figure 9: Degree of referent uncertainty, measured as $H(R|S,C)$, against Condition over blocks. Higher entropy values indicate a higher degree of ambiguity in context. The error bars represent bootstrapped 95% CIs. All conditions tend to decrease their referent uncertainty over successive blocks. This suggests the systems used by participants are moving toward an optimally simple configuration for identifying the intended referent in context.

There was also a significant effect of Blocks 2 and 3, indicating that average entropy across all 4 conditions decreased over repeated communication; however, this was moderated by marginally significant two-and three-way interactions between Blocks 2 and 3 and the experimental manipulations (Context-Type and Access to Context). The Shape-Different conditions (which start with low ambiguity) show relatively little decrease in ambiguity over blocks (as demonstrated by the effect sizes for Block 2 x Context-Type and Block 2 x Context-Type being of a similar magnitude to the simple effect of Block). Similarly, the significant interactions between Block and Access to Context suggest that the decrease in uncertainty is less pronounced in the Shared conditions.

	Estimate β	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.059	0.010	71.798	5.991	<0.001
Access to Context	-0.054	0.020	71.799	-2.744	0.008
Context-Type	-0.062	0.019	75.513	-3.194	0.002
Block 2	-0.030	0.007	4282.567	-4.507	<0.001
Block 3	-0.042	0.007	4284.909	-6.352	<0.001
Access to Context:Context-Type	0.045	0.039	75.529	1.150	0.254
Access to Context:Block 2	0.026	0.013	4282.553	1.974	0.048
Access to Context:Block 3	0.034	0.013	4284.886	2.548	0.011
Context-Type:Block 2	0.039	0.014	4282.548	2.874	0.004
Context-Type:Block 3	0.054	0.014	4285.762	3.994	<0.001
Access to Context:Context-Type:Block 2	-0.036	0.027	4282.538	-1.311	0.190
Access to Context:Context-Type:Block 3	-0.052	0.027	4285.744	-1.915	0.056

Table 4: Results for Referent Uncertainty in Context, measured as $H(R|S,C)$, with Access to Context, Context-Type and Block as fixed effects. The dependent variable is continuous.

4. Discussion

4.1. Does contextual predictability shape signal autonomy?

We put forward the general hypothesis that contextual predictability shapes the degree of signal autonomy. To test this claim we manipulated both the sender's ability to access to the context in which their utterances were interpreted and the variability of context-types across trials. When the context is predictable, senders organise languages to be less autonomous (more context-dependent), exploiting contextual information to reduce uncertainty about the intended referent with highly compressible semantic representations. For conditions with lower contextual predictability, senders use more autonomous signals, and rely less on contextual information to discriminate between possible referents. In line with previous work, these results demonstrate that languages adapt to contextual constraints by reaching a tradeoff between informativeness and simplicity (Piantadosi et al., 2012; Silvey et al., 2014; Winters et al., 2015; Tinits et al., 2017).

The key finding is that number of unique signals used and referent uncertainty are predicted by both Context-Type and Access to Context. Furthermore, even though referent uncertainty varies substantially between conditions, all of the communication systems which develop during interaction are communicatively functional (i.e., capable of discriminating between referents in context), as indexed by our measures of communicative accuracy and referent uncertainty in context. If the Context-Type is stable for discrimination across trials (Shape-Different), and speakers have access to this contextual information (Shared Context), then participants will produce higher out-of-context referent uncertainty (indicating non-autonomous signals) than participants in conditions where one or both of these variables is less predictable.

Having highly predictable contextual information allows senders to reach the most compressible set of signal-referent mappings capable of identifying the intended referent in context (Kirby et al., 2015). By using the minimal number of signals, and generalising across unique referents, receivers were able to extract information from both the signal and the context to learn what is and is not relevant for the task of discrimination. In the Shape-Different + Shared Context, which was maximally predictable in terms of context-type and access to context, this meant that communicative success remained constant and high across blocks: Senders used their knowledge of the context to leave out the colour-dimension, as this was irrelevant to communicative success, and only conveyed the shape-dimension in the linguistic system (see Figure 10 for example language).

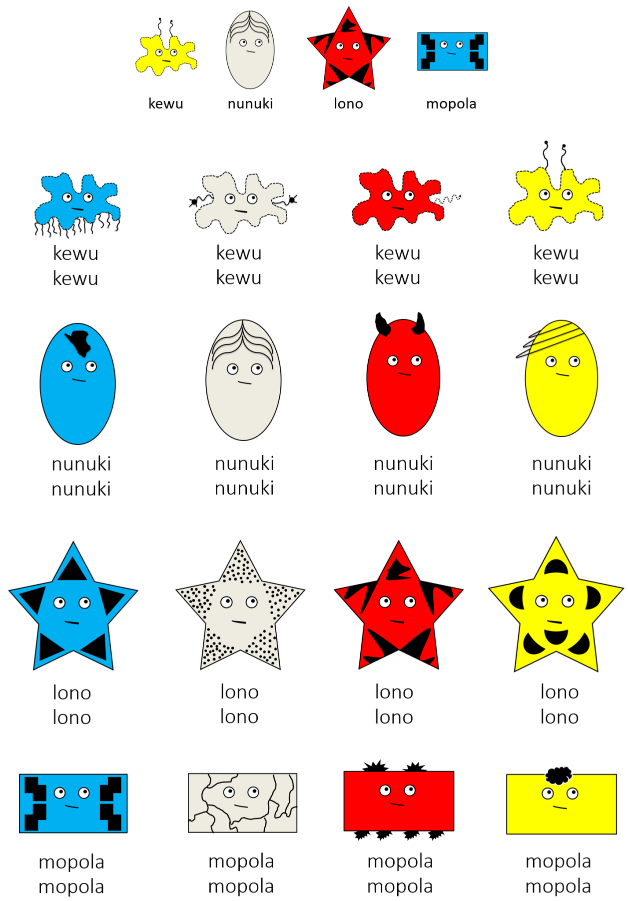


Figure 10: An example language from the final block of the Shape Different + Shared Context condition. The sender labelled each referent twice during interaction, both labels are shown here. In this case, the participants maintained the original four labels they were trained on (top row), generalising them to only encode information about the shape dimension.

As contextual predictability decreases, the sender is unable to estimate, and therefore exploit, this contextual information when designing their utterances. This inability to exploit contextual information results in an increased pressure to create autonomous signals (i.e., signals which are identifiable out of context). Autonomous systems are advantageous where contextual information is hard to predict, but they are also cognitively expensive; senders must remember, and receivers must learn, a greater number of conventional signals. For instance, in the Mixed + Unshared Context condition, which was maximally unpredictable in terms of context-type and access to context, senders specified both colour and shape within the signals (see Figure 11). This possibly explains the lower starting point and the gradual increase in communicative success score for receivers in Mixed + Unshared Context: Not only are receivers required to learn a larger set of conventional signals, they must also extract information from the internal structure of these signals.

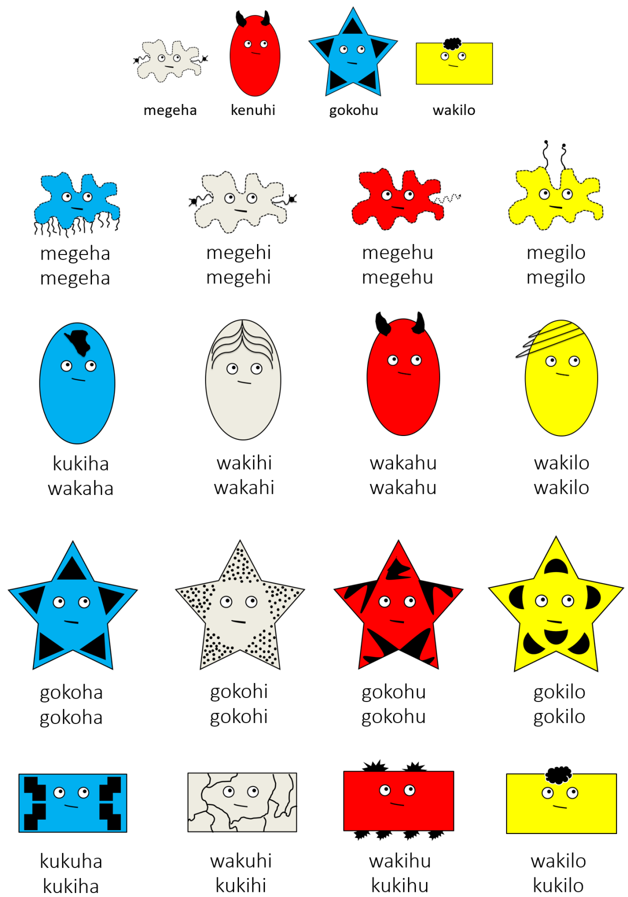
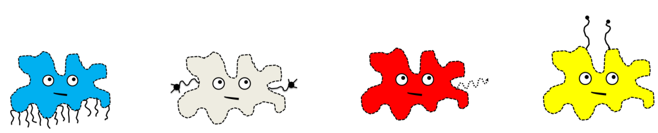


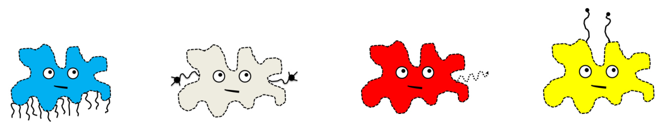
Figure 11: An example language from the final block of the Mixed + Unshared Context condition (top row with smaller images corresponds to the initial training language). Under each referent are two signals arranged according to trial order (i.e., the first signal was produced prior to the second signal for conveying a particular referent). Although there is some variability, the system can be generally described as having signals with two component parts: the initial component identifies shape and the final component identifies colour. For example, if the initial component is *mege*, then it refers to a blob, and if the final syllable is *ha* it refers to blue. The combination of these component parts results in the signal *megeha*, referent blue blob.

For participants in Shape-Different + Unshared Context, a gradual increase in communicative success is associated with a drop in the total number of unique signals, and an increase in out-of-context referent uncertainty: on average, senders end up using fewer signals to convey more referents, resulting in increasingly context-dependent systems (this is illustrated in Figure 12 where part of a signalling system becomes more context-dependent over successive blocks). The opposite is true for participants in Mixed + Shared Context: the average out-of-context referent uncertainty decreases as the communicative success increases, with the set of signal-referent mappings transitioning from a one-to-many to a one-to-one mapping (see Figure 13 for an example where part of a system becomes less context-dependent over successive blocks). The divergence between these two conditions suggests context-type exerts a stronger effect than access to context on the types of systems which emerge (as reflected in the reported effect sizes). Systems in Shape-Different + Unshared Context evolve to become increasingly context-dependent (and move closer to the simpler systems found in Shape-Different + Shared Context) whereas the systems in Mixed + Shared Context become more autonomous (and move closer to the relatively complex systems found in Mixed + Unshared Context).



block 1	kim kim	lunimugi kimunus	kimunus kinumu	hamoho kimugi
block 2	kim kim	kinumus kinumus	kinumu kinumus	kinumus kimunugi
block 3	kim kim	kinumus kinumus	kinumus kinumus	kinumus kinumus

Figure 12: A subset of a gradually emerging context-dependent system in the Shape-Different + Unshared Context condition. At the first block, the system is relatively autonomous, in that each blob has a unique signal that identifies whether it is a blue, grey, red, or yellow one. However, by block 3 the signal *kinumus* refers to the grey, red, and yellow blobs (the exception being the blue blob which retains a unique identifying signal, *kim*).



block 1	memuno memunopewa	memuno memunonoka	memuno memunomemuno	memuno memunokihimo
block 2	memunopewa memunopewa	memunonoka memunonoka	memuno memunomemuno	memunokihimo memunokihimo
block 3	memunopewa memunopewa	memunonoka memunonoka	memuno memunomemuno	memunokihimo memunokihimo

Figure 13: A subset of a language in the Mixed + Shared Context condition showing the emergence of an autonomous system. The first row in each block contains signals used in shape-different contexts and the second row contains signals used in colour-different contexts. At block 1, the system is context-dependent, with an underspecified signal, *memuno*, being used in shape-different contexts to refer to blob, and compositional signals, *memunopewa*, *memunonoka*, *memunomemuno*, and *memunokihimo*, being employed for colour-different contexts. By block 3 the underspecified mappings are no longer used and the system is no longer context-dependent (overspecified mappings are used in both shape-different and colour-different contexts). The only exception is for the red blob, which uses *memuno* in the shape-different context and a reduplicated form, *memunomemuno*, in the colour-different context.

4.2. Context-Type: Historical Context

The stronger effect of context-type points to the importance of historical relationships between contexts. Previous work showed that the preceding contexts, and the communicative solutions derived from these contexts, can shape the outcomes in the immediate context (Brennan & Clark, 1996; Barr & Keysar, 2002; Van Der Wege, 2009). We extend upon this work in two ways. First, our focus is on the emergence of a communication system, as opposed to looking at a natural language with a pre-established set of communicative strategies. Second, by directly manipulating what is locally relevant in the immediate context, we were able to investigate scenarios where solutions in the local context are either generalisable or not generalisable across the set of aggregate contexts.

For Shape-Different context conditions, what is locally relevant in the immediate referential context is also globally relevant across the total set of aggregated contexts (i.e., only encoding shape is relevant for communicative success). This increases the probability of senders discovering, and then persisting with, the most most compressible system capable of identifying the sender's intended referent in context. Mixed Context conditions, on the other hand, have a mismatch between what is locally relevant and what is globally relevant. A locally relevant solution in the colour-different context is to encode colour, but this solution does not generalise to the trials where the contexts are shape-different.

The simplest possible system capable of communication in Mixed + Shared Context would be one where the initial language is conditioned on the context-type. For example, using the initial language from Fig. 14, this consists of a system where *megeha* identifies blob-shaped images in Shape-Different contexts and grey coloured referents in Colour-Different contexts. Interestingly, no participants devised such a system. This difficulty in discovering a maximally simple system for communication explains the prevalence of autonomous systems in Mixed conditions. Encoding both shape and colour as distinct signal components is advantageous because it not only reduces uncertainty about the intended referent in the immediate context, but it also reduces future predictive uncertainty, i.e., a signal that encodes both colour and shape can be directly used in both shape-different and colour-different trials, rather than requiring senders or receivers to flip between two signals for a given referent depending on the context. Another advantage is that senders and receivers only have to rely on learning a set of conventional signals, as opposed to tracking variable context-types, in order to achieve their respective communicative goals. Whether or not this reduces the cognitive load for interlocutors remains an open question. Future work would be well-placed to investigate the different demands of learning conventions versus tracking contextual information.

Conventions also play an important role in terms of how senders generalise from their initial training language. We frequently noted that senders not only preserve the original signal-referent mappings they were trained on, they also use this common set of pre-established conventions as a basis for building the entire language. As an example, the signal *gokohu* in Figure 14 unambiguously refers to the blue star, irrespective of whether it was in a colour-different or shape-different context. However, when abstracted across the entire system, *gokohu* can be used to refer to any blue object as well as any star-shaped object. While this system exhibits ambiguity at the lexical level, complete utterances are interpretable out of context and therefore autonomous, as combinations of distinct signals unambiguously convey the intended referent. The fact that the training forms remain underspecified during communication shows that historical precedent is a powerful factor in determining outcomes (Brennan & Clark, 1996; Millikan, 1998).

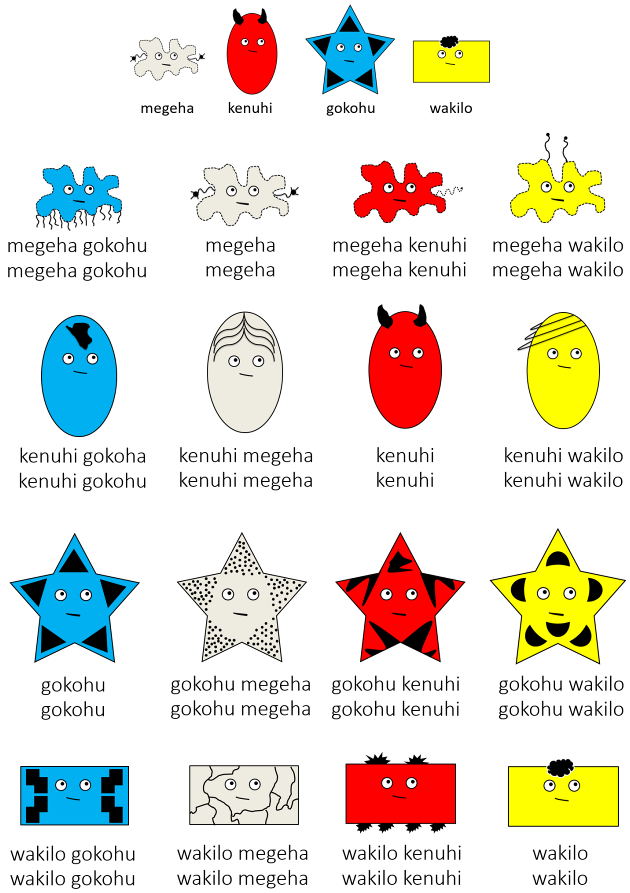


Figure 14: An example language from the final block of the Mixed + Shared Context condition (top row with smaller images corresponds to the initial training language). Under each referent are two signals: the first corresponds to the signal used in Shape-Different context-types and the second is the signal used in Colour-Different context-types. A single signal is composed of two parts: the initial signal refers to shape (e.g., *wakilo* refers to rectangle) and the second signal refers to colour (e.g., *wakilo* refers to yellow), regardless of context type. Also note that senders keep the ambiguous mappings they were trained on (e.g., *wakilo* for the yellow rectangle).

4.3. Shared Context

This is not to say that access to a shared context should be ignored. On the contrary, the amount of shared context between sender and receiver is relevant, with unshared contexts tending to have higher signal autonomy than shared contexts. For example, when compared with Shape-Different + Shared Context, senders in Shape-Different + Unshared Context often produce autonomous forms, and persist with these systems even though a simpler, more compressible system is sufficient. Not having access to the set of distractors means that senders can only indirectly infer the optimal solution via trial and error feedback. Whether or not senders did shift to an optimal system was contingent on their prior assumptions as well as how they modified the signal-referent mappings in response to feedback (see section 4.4).

Having a shared context also allowed senders, in some cases, to override the countervailing effects of a mixed context-type. This is because having access to the target and its distractors increases contextual predictability: senders can extract information about what is locally relevant in immediate context as well as what is globally relevant across the set of aggregated contexts. Take the two Mixed conditions: if you placed the systems found in these conditions on a cline of more to less autonomous, those found in Mixed + Shared Context can generally be described as less autonomous than those found in Mixed + Unshared Context. An example of this is where some systems in the Mixed + Shared Context are characterised by marked and unmarked forms depending on the context (see Figure 15): here, an unmarked signal (e.g., *kewu*) is used to convey shape information (e.g., blob) in shape-different contexts, whereas in colour-different contexts marked forms are used (e.g., *kewu mopola* for the referent blob blue).

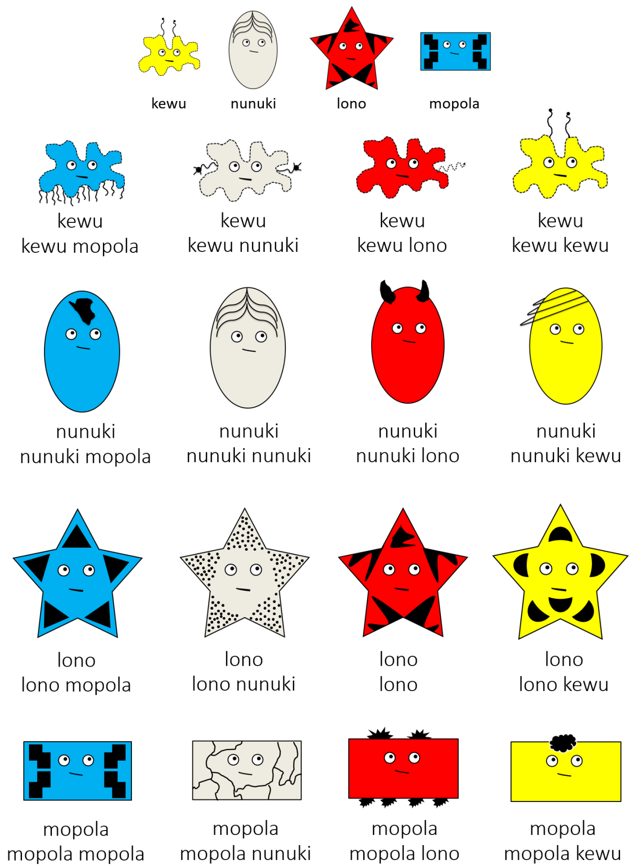


Figure 15: An example language from the final block of the Mixed Shared condition (top row with smaller images corresponds to the initial training language). Under each referent are two signals: the first corresponds to the signal used in Shape-Different context-types and the second is the signal used in Colour-Different context-types. Notice that in the Shape-Different context-types an underspecified signal is used, whereas in Colour-Different context-types an overspecified signal is used (where information about both colour and shape dimensions is encoded through the use of word order). For example, using the underspecified signal, *mopola*, in shape-different contexts refers to rectangle, whereas in colour-different contexts each referent has a unique overspecified signal (*mopola mopola*, *mopola nunuki*, *mopola lono*, *mopola kewu*). Interestingly, the overspecified system is order-dependent, with a word initial *mopola* referrring to rectangle, whereas a word final *mopola* refers to the colour blue.

4.4. Convergence and Divergence in Individual Behaviour

Discovering an optimal system is not solely contingent on context-type or the amount of shared context: prior assumptions of senders, as well as the extent to which they integrate information from feedback, are candidate factors for why a system might transition to a context-dependent or autonomous state.

First of all, there was considerable variation in whether the senders started out using autonomous or non-autonomous signals. This variation between individuals lends weight to the idea of populations being composed of heterogeneous pragmatic reasoners; speakers do not necessarily start out with the same initial assumptions, even if they eventually converge on the same set of behaviours (Franke & Degen, 2016). Our approach shows how to unmask these differences between individuals by manipulating whether contextual predictability is *reinforcing* (i.e., both manipulations either decrease or increase predictability) or *conflicting* (i.e., when one manipulation increases predictability and the other decreases predictability). When contextual predictability is reinforcing, as is the case in Shape-Different + Shared Context and Mixed + Unshared Context, participants are more likely to converge on similar systems. In cases where the contextual predictability is conflicting, as in Shape-Different + Unshared Context and Mixed + Shared Context, senders with differing initial assumptions can produce radically different systems of communication.

Another contributing factor to individual behaviour is how senders adjust their signals based on feedback. This is especially relevant for senders in both Unshared conditions: here, the only extractable information about the receiver's behaviour is whether or not they selected the target referent. By contrast, senders in the Shared conditions could integrate feedback with their prior knowledge of previous contexts and information about the immediate context. As such, the extent to which senders incorporate feedback in guiding their future communicative behaviour is an important question, and might be a contributing factor as to why some senders persisted with their original system and others decided to make their systems more or less autonomous.

If real-world languages are subject to similarly weak constraints, and individual variation plays an outsized role in determining the trajectory of change, then one general prediction is that cross-linguistic variation should not straightforwardly reflect a direct relationship between contextual predictability and signal autonomy. Instead, the outcomes will to some extent be historically contingent, albeit with the space of possible languages being bounded by cognitive, contextual, and communicative factors.

4.5. Future work

The current experiment was far from exhaustive and there are several features which which could be addressed in future work to improve its relevance to real-world linguistic interactions. The first of these is the use of artificial languages. As we were expressly concerned with the emergence of structure, the use of natural language was prohibited, but it might be the case that there are fundamental differences between the emergence of languages in this experiment and the use of languages with a set of pre-existing conventions. Similarly, the requirement for participants to not use their native language potentially increases the cognitive load, and this might impact the generalisability of these results. For instance, senders may optimise more in Mixed conditions when using their native language, as this removes the need to negotiate a set of conventional signals from scratch.

Second, knowledge of the context is relatively circumscribed in this experiment, whereas real-world language use draws on a much vaster set of relevant contexts. Under these situations, interlocutors might use heuristics to coordinate on only a small subset of information, ignoring other relevant factors (Gigerenzer & Gaissmaier, 2011). This aligns to some extent with the findings in the current experiment: senders tend to rely more on context-type in determining signal autonomy, minimising the effect of access to context. But it remains an open question as to whether or not introducing additional parameters, such as varying the the number of objects in the referential context (see Rubio-Fernández, 2016), enhances or diminishes the interactions between other variables associated with contextual predictability.

Lastly, in communication game experiments, we still know very little about the extent to which senders rely more on their private knowledge of the context Horton & Gerrig, 2005, as opposed to actively monitoring the needs of the receiver Clark & Carlson, 1981. All we assumed in this experiment was that the amount of shared context impacts signal autonomy by either decreasing or increasing contextual predictability. Future work could build on the present experiment to examine whether the presence or absence of a receiver impacts the behaviour of the sender. Another possibility is to provide more ecologically relevant conversational repair (e.g., Dingemanse et al., 2015) and turn-taking (Schegloff & Sacks, 1977) mechanisms. This would also allow us to test a broader range of communicative goals, besides our simple characterisation of communication as reducing uncertainty about the intended referent, such as the extent to which senders monitor and probe mutual understanding (Blokpoel et al., 2012).

5. Conclusion

A good system of communication must reach a balance between informativeness (to identify the sender's intended referent) and simplicity (to make the system compressible). We set out to investigate how these two pressures are influenced by the context in which languages are used. We showed that contextual predictability shapes the degree of signal autonomy. When the context is predictable, senders use this reliable information to devise compressible languages, whilst also maintaining the minimal requirement of being informative about their intended referent. This results in low autonomy: the signals in these systems are dependent on contextual information for disambiguation. Conversely, when the context decreases in predictability, senders increasingly rely on the internal structure of the signals to reduce uncertainty about the intended referent, resulting in a greater level of autonomy.

References

Arts, A., Maes, A., Noordman, L., & Jansen, C. (2011). Overspecification facilitates object identification. *Journal of Pragmatics*, *43*, 361–374.

Barr, D. J., & Keysar, B. (2002). Anchoring comprehension in linguistic precedents. *Journal of Memory and Language*, *46*, 391–418.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278.

Bates, D., Mchler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Usinglme4. *Journal of Statistical Software*, *67*.

Bazire, M., & Brézillon, P. (2005). Understanding context before using it. *Modeling and using context*, (pp. 113–192).

Beuls, K., & Steels, L. (2013). Agent-based models of strategies for the emergence and evolution of grammatical agreement. *PloS one*, *8*, e58960.

Blokpoel, M., van Kesteren, M., Stolk, A., Haselager, P., Toni, I., & Van Rooij, I. (2012). Recipient design in human communication: simple heuristics or perspective taking? *Frontiers in human neuroscience*, *6*.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1482.

Brochhagen, T., Franke, M., & van Rooij, R. (2016). Learning biases may prevent lexicalization of pragmatic inferences: a case study combining iterated (bayesian) learning and functional selection. In *Proceedings of the 38th annual conference of the cognitive science society* (pp. 2081–2086).

Chater, N., & Vitányi, P. (2003). Simplicity: a unifying principle in cognitive science? *Trends in cognitive sciences*, *7*, 19–22.

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

Clark, H. H., & Carlson, T. B. (1981). Context for comprehension. *Attention and performance IX*, (pp. 313–330).

Cover, T. M., & Thomas, J. A. (1991). *Elements of Information Theory*. John Wiley & Sons Inc.

Croft, W. (2000). *Explaining language change: An evolutionary approach*. Pearson Education.

Culbertson, J., & Kirby, S. (2016). Simplicity and Specificity in Language: Domain-General Biases Have Domain-Specific Effects. *Frontiers in Psychology*, *6*.

Dale, R., & Reiter, E. (1995). Computational interpretations of the gricean maxims in the generation of referring expressions. *Cognitive science*, *19*, 233–263.

Dingemanse, M., Roberts, S. G., Baranova, J., Blythe, J., Drew, P., Floyd, S., Gisladottir, R. S., Kendrick, K. H., Levinson, S. C., Manrique, E. et al. (2015). Universal principles in the repair of communication problems. *PloS one*, *10*, e0136100.

Duranti, A., & Goodwin, C. (1992). *Rethinking context: Language as an interactive phenomenon* volume 11. Cambridge University Press.

Evans, V., & Green, M. (2006). *Cognitive linguistics. An introduction*. Edinburgh University Press, Edinburgh.

Frank, M. C., & Goodman, N. D. (2012). Predicting Pragmatic Reasoning in Language Games. *Science*, *336*, 998–998.

Franke, M., & Degen, J. (2016). Reasoning in Reference Games: Individual- vs. Population-Level Probabilistic Modeling. *PLOS ONE*, *11*, e0154854.

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

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual review of psychology*, *62*, 451–482.

Gil, D. (2005). From repetition to reduplication in riau indonesian. *Studies on reduplication*, *28*, 31–64.

Grodner, D., & Sedivy, J. C. (2011). 10 the effect of speaker-specific information on pragmatic inferences. *The processing and acquisition of reference*, (p. 239).

Heller, D., Skovbroten, K., & Tanenhaus, M. K. (2009). Experimental evidence for speakers sensitivity to common vs. privileged ground in the production of names. *Proc. PRE-CogSci*, .

Horton, W. S., & Gerrig, R. J. (2005). The impact of memory demands on audience design during language production. *Cognition*, *96*, 127–142.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, *59*, 91–117.

Hurford, J. R. (2011). *The origins of grammar: Language in the light of evolution II* volume 2. Oxford University Press.

Isaacs, E. A., & Clark, H. H. (1987). References in conversation between experts and novices. *Journal of experimental psychology: general*, *116*, 26.

Kay, P. (1977). Language evolution and speech style. In *Sociocultural dimensions of language change* (pp. 21–33). Elsevier.

Kemp, C., & Regier, T. (2012). Kinship categories across languages reflect general communicative principles. *Science*, *336*, 1049–1054.

Keysar, B., Barr, D. J., & Horton, W. S. (1998). The egocentric basis of language use: Insights from a processing approach. *Current directions in psychological science*, *7*, 46–49.

Kirby, S. (1999). *Function, selection, and innateness: The emergence of language universals*. OUP Oxford.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102.

Konopka, A. E., & Brown-Schmidt, S. (2014). Message encoding. *The Oxford handbook of language production*, (pp. 3–20).

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). lmertest: tests in linear mixed effects models. r package version 2.0-20. *URL: https://cran. rproject. org/web/packages/lmerTest. Accessed*, *15*, 2016.

Lewis, D. (1969). Convention cambridge. *Mass.: Harvard UP*, .

Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PloS one*, *5*, e8559.

Millikan, R. G. (1998). Language conventions made simple. *The Journal of Philosophy*, *95*, 161–180.

Moreno, M., & Baggio, G. (2015). Role asymmetry and code transmission in signaling games: an experimental and computational investigation. *Cogn Sci*, *39*, 918–43.

Nadig, A. S., & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, *13*, 329–336.

Olson, D. R. (1970). Language and thought: Aspects of a cognitive theory of semantics. *Psychological review*, *77*, 257.

Pechmann, T. (1989). Incremental speech production and referential overspecification. *Linguistics*, *27*, 89–110.

Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition*, *122*, 280–291.

Ramscar, M., & Port, R. (2015). Categorization (without categories). *Handbook of Cognitive Linguistics. De Gruyter Mouton*, .

Regier, T., Kemp, C., & Kay, P. (2015). Word Meanings across Languages Support Efficient Communication. In *The Handbook of Language Emergence* (pp. 237–263). John Wiley & Sons Inc.

Rubio-Fernández, P. (2016). How redundant are redundant color adjectives? an efficiency-based analysis of color overspecification. *Frontiers in psychology*, *7*.

Schegloff, J. G., E.A., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, *53*, 361–382.

Scott-Phillips, T. C. (2015). *Speaking Our Minds*. Macmillan Education UK.

Sedivy, J. C. (2005). 17 evaluating explanations for referential context effects: Evidence for gricean mechanisms in online language interpretation. *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions*, (p. 345).

Silvey, C., Kirby, S., & Smith, K. (2014). Word Meanings Evolve to Selectively Preserve Distinctions on Salient Dimensions. *Cognitive Science*, *39*, 212–226.

Snow, C. E., Barnes, W. S., Chandler, J., Goodman, I. F., & Hemphill, L. (1991). *Unfulfilled expectations: Home and school influences on literacy.*. Harvard University Press.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition*. Cambridge, MA, USA: Harvard University Press.

Steels, L. (2003). Evolving grounded communication for robots. *Trends in Cognitive Sciences*, *7*, 308–312.

Tamariz, M. (2017). Experimental Studies on the Cultural Evolution of Language. *Annual Review of Linguistics*, *3*, 389–407.

Terkourafi, M. (2009). On de-limiting context. *Contexts and constructions*, *9*, 17.

Tinits, P., Nölle, J., & Hartmann, S. (2017). Usage context influences the evolution of overspecification in iterated learning. *Journal of Language Evolution*, .

Trudgill, P. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford University Press.

Van Der Wege, M. M. (2009). Lexical entrainment and lexical differentiation in reference phrase choice. *Journal of Memory and Language*, *60*, 448–463.

Winters, J., Kirby, S., & Smith, K. (2015). Languages adapt to their contextual niche. *Language and Cognition*, *7*, 415449.

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.

Yoon, S. O., Benjamin, A. S., & Brown-Schmidt, S. (2016). The historical context in conversation: Lexical differentiation and memory for the discourse history. *Cognition*, *154*, 102–117.

Zifp, G. (1949). *Human Behavior and the Principle of Least Effort*. Addison-Wesley, New York.