Cognition 182 (2019) 151–164

Contents lists available at ScienceDirect

Cognition

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

Original Articles

Compositional structure can emerge without generational transmission

Limor Raviv[a,*], Antje Meyer[a,b], Shiri Lev-Ari[a,c]

[a] Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands
[b] Radboud University Nijmegen, the Netherlands
[c] Royal Holloway University of London, Egham, UK

ARTICLE INFO

Keywords:
Language evolution
Iterated learning
Communication
Input variability
Artificial language experiments
Compositionality

ABSTRACT

Experimental work in the field of language evolution has shown that novel signal systems become more structured over time. In a recent paper, Kirby, Tamariz, Cornish, and Smith (2015) argued that compositional languages can emerge only when languages are transmitted across multiple generations. In the current paper, we show that compositional languages can emerge in a closed community within a single generation. We conducted a communication experiment in which we tested the emergence of linguistic structure in different micro-societies of four participants, who interacted in alternating dyads using an artificial language to refer to novel meanings. Importantly, the communication included two real-world aspects of language acquisition and use, which introduce compressibility pressures: (a) multiple interaction partners and (b) an expanding meaning space. Our results show that languages become significantly more structured over time, with participants converging on shared, stable, and compositional lexicons. These findings indicate that new learners are not necessary for the formation of linguistic structure within a community, and have implications for related fields such as developing sign languages and creoles.

1. Introduction

Amongst the most important questions in the field of language evolution are how and why linguistic structure emerged, and under which pressures it evolved (Bickerton, 2007). According to usage-based theories, language is an adaptive and culturally transmitted system that has evolved to fit speakers' cognitive biases and constraints (Deacon, 1997; Reali & Griffiths, 2009; Smith, 2011) and to maximize their communicative success (Beckner et al., 2009; Mirolli & Parisi, 2008). A critical phase in the process of language evolution is the transition from an unstructured proto-language to a state of a full-blown language that exhibits compositional structure (Jackendoff, 1999; Zlatev, 2008). Compositionality, i.e., the systematic recombination of small units to express different meanings, is considered one of the hallmarks of natural language, which differentiate it from animal communication systems (Hockett, 1960). Indeed, one of the things that makes natural languages so unique is their infinite expressive power, which is the direct result of compositionality: we can talk about an unlimited set of meanings thanks to our ability to recombine a limited set of sub-elements in systematic ways.

In the past two decades, two different strands of experimental work have attempted to investigate the factors involved in the emergence of linguistic systems from two distinct perspectives. First, Experimental Semiotics studies focused on the communicative and social nature of language evolution, and examined how interactions between pairs or groups influence convergence, iconicity and complexity of visual signals (e.g., Galantucci & Garrod, 2011; Garrod, Fay, Lee, Oberlander, & MacLeod, 2007). In Experimental Semiotics studies, the main pressure is a communicative pressure for expressivity: signals should be expressive, informative and communicatively efficient in order to allow for reliable discrimination between potential referents, and should be shared across participants to allow for mutual understanding. Second, Iterated Learning studies focused on how individuals' cognitive biases and constraints shape previously established signs over the repeated transmission to new generations of learners, and examined how signal systems change in terms of learnability and structure (e.g., Beckner, Pierrehumbert, & Hay, 2017; Kirby, Cornish, & Smith, 2008). In Iterated Learning studies, the main pressure is a learning pressure for compressibility: limitations on memory create a pressure for signals to become simpler, more compressed and more predictable, so that languages could be easily learned from a finite set of exemplars, and generalizable to a new set of exemplars (Kirby, Griffiths, & Smith, 2014; Kirby et al., 2008). Both these literatures have generated numerous novel findings with important implications for the evolution of language. For example, Experimental Semiotics paradigms have been used to examine the emergence of arbitrary signals from iconic signs (e.g.,

* Corresponding author at: Max Plank Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, the Netherlands.
E-mail address: limor.raviv@mpi.nl (L. Raviv).

https://doi.org/10.1016/j.cognition.2018.09.010
Received 25 September 2017; Received in revised form 13 September 2018; Accepted 13 September 2018

Garrod et al., 2007). Iterated Learning has typically been used to examine the creation of compositional regularities (e.g., Kirby et al., 2008), but has also been used to examine the evolution of case markers (e.g., Smith & Wonnacott, 2010) and color terms (e.g., Xu, Dowman, & Griffiths, 2013).

In a recent and highly influential study, Kirby, Tamariz, Cornish, and Smith (2015) combined the paradigms of Experimental Semiotics and Iterated Learning and contrasted two experimental conditions: communication with transmission vs. communication without transmission. In the communication and transmission condition (the "chain" condition), pairs of participants communicated about a structured meaning space using an artificial language, and then their languages were transmitted to new pairs of participants over several generations. In the communication without transmission condition (the "closed group" condition), pairs interacted amongst themselves for several rounds, with no new learners being introduced over time. The results showed that when languages were transmitted over multiple generations of pairs, they developed compositional, morphology-like structures in which different affixes were systematically combined to express similarities in meanings. In contrast, when the same pairs communicated for repeated rounds without generational turnover, they created holistic, unstructured languages in which each item was assigned a unique label, and feature overlap between items was not reflected in the labels.

Kirby et al. (2015) argued that the reason compositionality did not emerge in the closed-group condition is because pairs were able to get highly familiar with the signs, so there was no reason for them to develop compressed, systematic structures instead of holistic languages. They interpret their results as showing that (1) compositionality arises only as a tradeoff between expressivity and compressibility pressures; and (2) that expressivity and compressibility pressures stem from two independent sources - communication and transmission – which operate at different timescales. Kirby et al. (2015) view these two processes as bringing about conflicting constraints: while horizontal intra-generational communication pushes languages to become maximally expressive, vertical cross-generational transmission pushes languages to become maximally compressed. By providing a systematic mapping between meanings and signals, compositionality offers an equilibrium between the need to minimize the associated memory and cognitive costs while maximizing languages' expressivity. This idea suggests that the basic architecture of natural language can be explained by the interaction of conflicting weak cognitive biases and processing limitations, and by taking the pragmatic context in which languages evolve into account (Christiansen & Chater, 2016; Culbertson & Kirby, 2016).

Importantly, Kirby et al. (2015) fully equate expressivity and compressibility pressures with communication and transmission respectively. They argue that horizontal communication gives rise to expressivity pressures due to people's communicative goals: languages should be expressive given the need to interact and successfully discriminate between different meanings. Vertical transmission is argued to give rise to compressibility pressures due to people's memory limitations and cognitive biases: languages should be simple and easy to learn given that are being repeatedly learned over generations by new people. They predict that compositionality emerges only when both communication and transmission are at play, as a solution to these competing pressures. On one hand, a compressibility pressure operating in isolation (e.g., languages are only transmitted across generations of learners, but not used for communication) leads to underspecified languages with minimalistic lexicons, where multiple meanings are represented with a single word (as found in Kirby et al., 2008). While such simple systems are highly compressed and easy to learn, they are degenerated, ambiguous and lacked expressivity. On the other hand, an expressivity pressure operating in isolation (e.g., languages are only used for communication, but never transmitted to new learners) should potentially result in languages with massive lexicons, where each meaning is represented with a unique word. While such holistic systems would be maximally expressive, they would also be incompressible and therefore hard to learn and remember by new individuals. If languages need to be both expressive and compressed (i.e., because they are being used for communication as well as being transmitted to new learners), developing regularities in the form of compositional structure will maintain their informativity while reducing the memory load and increasing languages' learnability. This is because compositional languages allow for the expression of multiple different meanings using a recombination of the same basic elements. As such, a compositional language is highly compressed and simpler in comparison to a holistic language (where the same set of meanings would require memorizing more unique words), while also being highly expressive and informative in comparison to a degenerated language (where the same set of meanings would be indistinguishable). In sum, Kirby et al. (2015) predict that both communication and transmission are necessary for the emergence of compositionality, and conclude that communication alone (i.e., without generation turnover) is not enough for compositionality to emerge. This finding has since been replicated with different meaning spaces (Carr, Smith, Cornish, & Kirby, 2017; Winters, Kirby, & Smith, 2015) and with artificial sign languages (Motamedi, Schouwstra, Smith, & Kirby, 2016).

This conclusion has far-reaching implications for the literature on the evolution of language, as well as for the broader field of cultural evolution. First, it directly relates to work on creolization and emerging sign language by suggesting that one of the "design features" of natural language may need several generations to emerge. Supporting this idea, studies on the developing Nicaraguan sign language have shown that complex linguistic structure emerges over multiple cohorts of learners (Senghas, Kita, & Ozyurek, 2004), and work on pidgins has suggested that new child learners are required in order to develop recursion (Bickerton, 1984). Second, it affects the reasoning and predictions made about the structure of human lexicons over time: from understanding trends in metaphorical mappings (Xu, Malt, & Srinivasan, 2017) to measuring the entropy and informativity of words (Bentz, Alikaniotis, Cysouw, & Ferrer-i-Cancho, 2017). Going beyond language evolution and change, this conclusion has already influenced work on a wide range of human behaviors. For example, compressibility pressures during cross-generational transmission have been implied to play a role in explaining cross-cultural differences in folk tale complexity (Acerbi, Kendal, & Tehrani, 2017), musical universals (Trehub, 2015), and the propagation and stabilization of behavioral conventions (Scott-Phillips, 2017).

In the current paper we suggest that communication in the real world includes not only expressivity pressures, but also several sources for compressibility pressures. In other words, while we agree with Kirby et al. (2015) that both expressivity and compressibility pressures are necessary for the emergence of compositionality, we believe that both pressures are already present during real-world communication. Therefore, we predict that in contrast to Kirby et al.'s (2015) conclusion, compositionality can emerge during communication in a closed group without generational transmission. This prediction is in line with several non-linguistic communication studies, which found that compositional structure can emerge in signal systems during interaction alone. First, Selten and Warglien (2007) found that when pairs of