

The Impact of Communicative Constraints on the Emergence of a Graphical Communication System

Till Bergmann (tbergmann@ucmerced.edu)

Cognitive and Information Sciences, University of California, Merced, Merced, CA 95343

Rick Dale (rdale@ucmerced.edu)

Gary Lupyan (lupyan@wisc.edu)
Department of Psychology, University of Wisconsin-Madison, Madison, WI 53706

Abstract

We investigated the behavior of participants tasked with communicating in a novel environment. Participants had to use their mouse to draw graphical representations (termed squiggles in the game) of human faces in order to communicate with fellow players. Experiment 1 investigated the effect of varying features of the input images on the resulting drawings. Experiment 2 introduced varying comprehension conditions that were predicted to produce differences in how features of faces would be graphically represented. In experiment 1, the features of the different faces significantly shaped the structure of the resulting squiggles. In experiment 2, the structure of the squiggles was influenced by the environment in which they were interpreted.

Keywords: social interaction; language evolution; evolution of communication; social feedback; reference; games

Introduction

For an individual learner, language can be described as a moving target (Christiansen & Chater, 2008; Chater, Reali, & Christiansen, 2009). Languages are commonly assumed to undergo changes at various levels. Recent studies have sought to identify the extra-linguistic factors that may determine the direction of these changes. These perspectives articulate an adaptive approach to our communication system: Any existing variability in a population of language users may have their language move in particular directions because some of those variants are more useful or successful than others at propagating. For example, recent studies have argued that the social environment of languages is a very important factor for these changes and developments (Lupyan & Dale, 2010; Wray & Grace, 2007). In addition, humans are able to invent new communication systems on the fly, with the emergence of Nicaraguan Sign Language and Al-Sayyid Bedouin Sign Language being two of the most impressive cases (Meir et al., 2010).

In recent years, a number of studies have been conducted with a focus on the structure of new communication systems that emerge in a given environment. In these studies, participants have to use novel symbols to successfully communicate with other participants. Over time, participants achieve more efficiency in their communication by aligning their symbols (Fay et al., 2010; Galantucci, 2005; Garrod et al., 2007).

In another line of research, the same question has been tackled by the use of different “generations” of participants. The first generation of participants learns simple word-object patterns, and then passes this information on to a new generation of learners. After a couple of generations, struc-

ture emerges that was absent in the original language system (Kirby, Cornish, & Smith, 2008).

This literature offers new insights, and importantly new techniques, for exploring artificial communication systems as a testbed for the forces that may drive communicative change. The present paper aims to shed light on the ways structure emerges in a new communication system by combining these two experimental frameworks in a large-scale interactive communication game. Specifically, we look to how the structure of a referential environment — the space of things that a community is going to talk about, and how these things are encountered — shapes the manner in which individuals in that community communicate. A first, exploratory experiment investigated the different strategies participants would employ to communicate a given image and how this given image influences the structure of the produced symbol. A further experiment introduced different environmental conditions that affected the way symbols were structured.

Previous Work

The current study expands on previous work by Dale and Lupyan (2010), in which players connect via the Internet to a gaming platform, where they had to create and comprehend visual signs for objects in order to communicate with other participants. Successful communication was achieved when players were able to match a previously drawn sign to its corresponding object.

The experiment consisted of two trial types: the production (*speaking*) and comprehension (*listening*) trials. In a speaking trial, participants were shown an object, and after a short delay, had three seconds to “squiggle” an image by clicking and dragging their computer mouse on a provided canvas. Such a method is akin to putting a temporal bandwidth on the ability of the user to sketch the image. This was designed to avoid their using English orthography or detailed sketches. Participants had to focus on essential aspects of an image to relay it. A different approach was used by Galantucci (2005), who used a rotating tablet to keep subjects from using orthography. Because this game framework is used online, the temporal bandwidth is a straightforward way of doing this, while also moving the game along more quickly.

In a listening trial, players were shown a previously drawn squiggle, along with two pictures of objects. The participant then had to choose the matching picture by clicking on it. Feedback is given whether they chose the correct picture. In the original work, squiggles in the listening trial were cho-

sen according to an evolutionary algorithm factoring in the novelty of the squiggle and its previous comprehensibility. Squiggles that performed badly in listening trials thus “die out” and are no longer used.

Data from 60 users were collected, which resulted in around 1,400 produced squiggles and around 4,100 listening trials. Feedback by the players was generally very positive and many found it to be entertaining and rather addictive.

Results from this study supported the findings from other studies: *a*) Squiggles get simplified. The average size of a squiggle shrinks over gameplay; *b*) The evolutionary algorithm produced stability for most squiggles, despite opportunities for novel squiggles to replace them during listen trials and *c*) Squiggles were at first drawn highly iconically, but gradually evolved towards more simplified symbols as participants continued playing.

In the current paper, we utilize this framework as a testbed for exploring how environments shape the squiggles as participants play. The central hypothesis guiding the study was that users’ squiggles are shaped by their “listening” environments (i.e., distribution over the foil-target pairs). In what follows, we first describe our redesign of the system, and an initial exploration (Experiment 1) of how aspects of objects can predict aspects of squiggles. Following this, we describe a comparison of two conditions (Experiment 2) meant to more subtly explore the role of the distribution of listening trials in determining squiggle composition.

Experiment 1

Method

Game Design The approach in Dale and Lupyan (2010) was replicated in the present study. However, the game was re-written to take into account new technological developments to ensure maximal compatibility with current browsers. We chose a specific domain of objects for squiggles to make “reference” to: human faces. While the previous study used a non-standardized set of pictures, all pictures were now taken from the Face Database (Minear & Park, 2004). This had the advantage that all picture used the same light conditions, background, facial expressions and dimensions. Variability on those parts was thus excluded. 120 pictures were chosen with respect to their age and gender, using a 2x2 matrix, with 30 pictures in each set (cf. Table 1). In order to reduce the number of salient features in these images, pictures with very distinguishing features such as birthmarks and earrings were not chosen.

Table 1: Summary statistics for the different picture sets

Group	M (Age)	SD (Age)	Range (Age)
Young Males	21.35	2.35	18–28
Young Females	21.53	2.85	18–28
Old Males	73.37	7.47	61–91
Old Females	73.80	6.69	61–85

Participants connected to the gaming platform via the Internet. Players were instructed to squiggle the displayed images so that other players are able to match a squiggle to the corresponding image later.

In the speaking trial, one of the 120 pictures in the data set was randomly shown to the participant along with a framed “canvas” on which participants could draw (Fig. 1). Participants had five seconds (without delay) to draw on the canvas before a listening trial appeared. The time limit was increased from the earlier study as the delay until the countdown started was removed. The remaining time for the speaking trial was shown to the user below the canvas.

In the listening trial, a squiggle previously produced by a fellow player is animated on a canvas using the original speed and directions of the user who had drawn the squiggle. The matching image was then placed either above or below this squiggle, with a competitor image taking up the other position (Fig. 2). Players had to click on the matching image in order to successfully communicate, and received visual feedback whether they had made the right choice. The ratio of speaking to listening trials was 1:3.

An important constraint on communication was time – participants were given only 5 seconds to create the squiggle.

Remember: Draw a Squiggle representing the person you see so that other people can make the connection later!

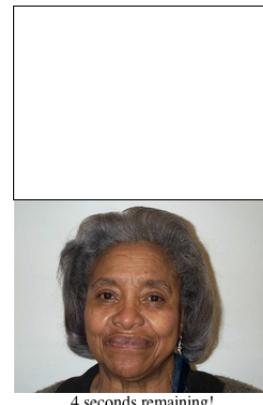


Figure 1: An example of a speaking trial.

Expectations Experiment 1 mainly served as a first step in testing the impact that different image dimensions have on the behavior of participants. It also served as an exploratory study of this behavior. We were interested in the different strategies employed by players to communicate. Which salient features were chosen and which were the most efficient? The input images were believed to have an effect on the way squiggles were structured by the participants; that is, we expected dimensions such as age and gender to influence squiggle strategies.

It was predicted that users would perform better in the listening trials with more practice, and hypothesized that the

Remember: Click the picture you think the Squiggle is representing!



Figure 2: An example of a listening trial.

complexity of produced squiggles would shrink over time. Players were expected to become more efficient with more exposure to the communication system in both the listening and speaking trials.

Participants 100 participants from the United States were recruited via Amazon Mechanical Turk and compensated financially (\$0.75) for taking part in the study. The use of Amazon Mechanical Turk had the indispensable advantage that players signed up almost instantaneously and simultaneously, which created a community of speakers playing at the same time and communicating with each other. Participants had to produce thirty squiggles and the respective number of listening trials (90) in order to receive compensation.

Data

On speaking trials, measurements of coordinates were taken along the paths drawn by the players. Time stamps were collected at each sampled x, y coordinate. The length in pixels of each squiggle was calculated using the stored coordinates. The number of unique strokes was collected as well.

In the listening trial, unique identifiers for the player, squiggle and corresponding image as well as the competitor image were stored. The comprehension of the result — whether the correct picture was chosen — was stored in a binary fashion along with the time it took the player to decide.

Data by participants who produced fewer than thirty squiggles were excluded, leaving data from 96 participants.

Results

Communication Strategies (Speaking) We performed a linear mixed-effects model on the relationship between the total length of a squiggle in pixels and the interaction of age, gender and progression of gameplay. Random effects consisted of the subject and item. p -values were obtained by calculating the likelihood of the corresponding t -values in a normal distribution. We found main effects for gender and age of the target, as well as for the progression of gameplay. Additionally, there was an effect for the interaction between age and gender. Squiggle length decreased for male input images ($t = -9.47, p < 0.0001$) by $129.6 \text{ pixels} \pm 13.68$ (standard error). Squiggle length decreased for young input images ($t = -3.19, p = 0.001$) by $43.33 \text{ pixels} \pm 13.68$. Length increased slightly over time by $1.1 \text{ pixels} \pm 0.39$ ($t = 2.71, p < 0.0001$). The interaction of age and gender meant that squiggles for young males were $42.24 \text{ pixels} \pm 19.37$ longer than for old males ($t = 3.91, p < 0.0001$), negating the main effect of age for males (Fig. 3).

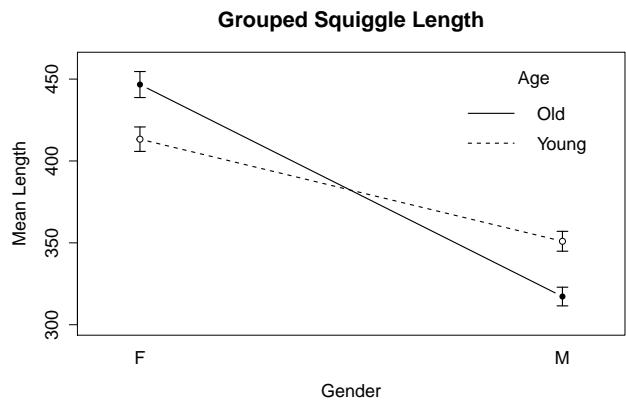


Figure 3: Average length of squiggles grouped by item age and gender

To account for possible differences in compositionality, we carried out a linear mixed-effects model on the relationship between the number of strokes and the interaction of age, gender and progressing gameplay. p -values were obtained by calculating the likelihood of the corresponding t -values in a normal distribution. We obtained main effects for age, gender and progressing gameplay. Squiggles contained more features ($+0.195 \pm 0.08$) if depicting a male ($t = -2.42, p = 0.015$) or young person ($+0.38 \pm 0.08, t = 4.80, p < 0.0001$). The number of strokes also increased slightly by 0.008 ± 0.002 with more practice ($t = 3.43, p < 0.001$). An effect of the interaction between age and gender ($t = -3.55, p = 0.0004$) meant that the main effect of age was negated for males, as old males elicited slightly more strokes (0.02) than young males (Fig. 4).

Comprehension Performance (Listening) Overall, participants chose the correct image 71% of the time, demonstrating that successful communication in novel environments is

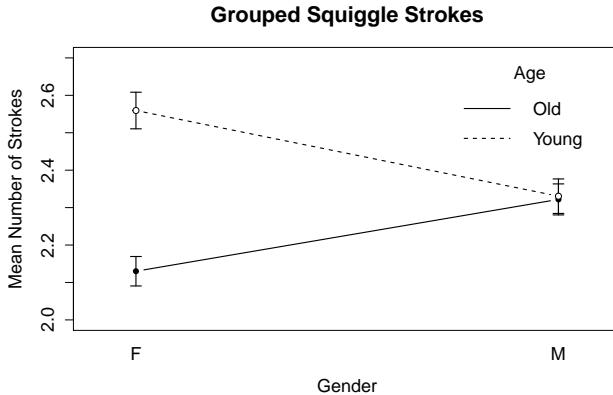


Figure 4: Average number of strokes grouped by item age and gender.

achievable even in a short period of time.

A linear-mixed effects model on the relationship of comprehension and the interaction of age, gender and progressing gameplay was constructed. An additional fixed effect was the time it took participants to make their decision. This variable was log-transformed to reform skew. Random effects were the users, items and the competitor images. The analysis revealed no significant effect for age, gender and progressing gameplay. The model produced a main effect for decision time ($z = -3.56, p = 0.0004$). Scores decreased by 0.18 \pm 0.05 the longer users contemplated their decision.

Discussion

As predicted, squiggles varied in structure and appearance depending on the image drawn. For females, participants focused on the hair form and length, which resulted in longer squiggles for females than males. Players tended to encode several, short features for younger females and fewer, but longer, features for old females. The increased number of strokes for young females can be explained by a preference for drawing hair which required several strokes (cf. Fig. 5a). Both strategies performed with the same level of accuracy in listening trials, which shows that each strategy chose the right aspects to be encoded.

Salient features of young males were often hair or head form, and on average, squiggles contained more features for males than for females. Old males were typically depicted by focusing on wrinkles and mouth form, which resulted in very short squiggles containing several features. The number of strokes did not vary much across age for males, as in both categories several features were encoded. A range of strategies across ages and genders is shown in figure 5. These were successful strategies, in that the squiggles were easy to be recognized among those who “listened” to it.

Both the number of encoded features and the total length of squiggles increased over time. This does not support our hypothesis that the complexity of squiggles would shrink over time. There are two explanations for this which are not mu-

tually exclusively. First, participants did not play the game long enough to get an effect. Second, there is no explicit pressure on the participants to reduce complexity. The time limit given in the speaking trial is sufficient enough to encode several features, and there is no motivation for users to change their strategy.

The fact that players have ample time in the speaking trial to produce detailed squiggles without much practice means that squiggles are comprehended right away. This leads to a ceiling effect and participants do not get better at comprehending squiggles over time. A decrease of the time limit in the speaking trial will very likely produce a decrease of comprehensibility at the beginning of the game, as players will not be able to produce detailed enough squiggles. Accuracy of comprehension should, however, increase over time in this scenario.

This initial study simply aimed at identifying basic strategies of players, observing the dynamics of gameplay (effect of time on production/comprehension), and whether squiggles can be quantified as having particular properties associated with particular classes of stimuli.

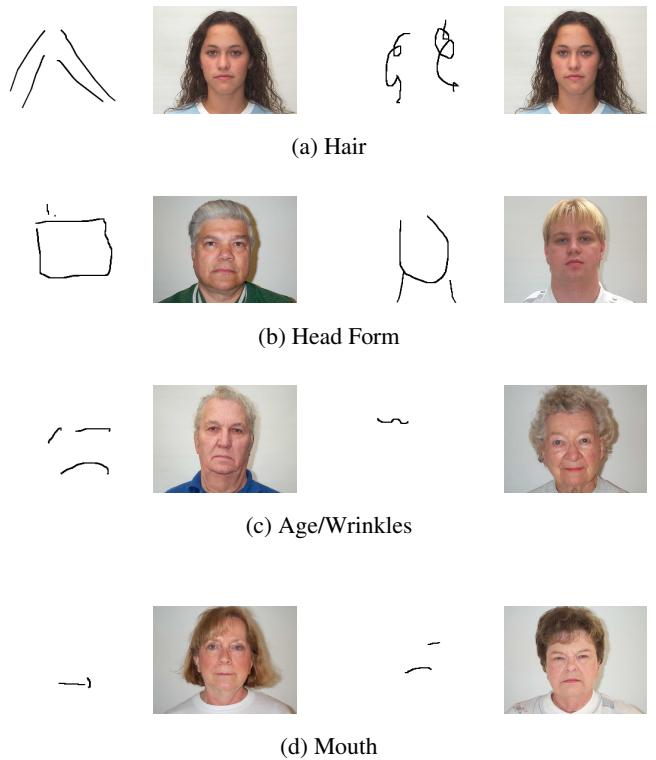


Figure 5: Different strategies employed in successful squiggles

Experiment 2

Method

Game Design After the initial experiment, a second, more refined experiment was carried out. We introduced a second condition in the listening trial, in which the competitor of the

displayed squiggle was always of the opposite age and gender. While age is normally a continuous variable, we chose the pictures in such a way that made a binary distinction possible (see above). For example, if the squiggle depicted a young male, the competitor would now always be an old female. The new condition provides additional contextual clues to the players: In order to match a squiggle, only age and gender need to be discriminated against instead of focusing on individual features. The other condition remained as in experiment 1, i.e. the competitor was randomly chosen with no regard to age or gender. A simple encoding of age and gender is not be sufficient in this condition.

The ratio of speaking trials to listening trials was changed additionally in both conditions. The higher rate of language comprehension in the early stages of language acquisition was modeled by increasing the number of listening trials to 10 for the first three speaking trials, and then slowly decreasing the number back to 3. Participants had to go through 20 speaking trials and 110 listening trials in total. No further changes were made to the game setup.

Expectations We predicted that the a changed comprehension environment would result in subsequent changes in the production of squiggles. In particular, we hypothesized that squiggles in the new condition would require less detail and complexity in order to achieve successful communication than squiggles in the control condition. The difference should be clearly visible after a short acquisition time. We expected a difference to occur roughly after the third speaking trial after which the number of listening trials was slowly decreased.

The expected differences in the structure of squiggles were predicted to not impact the performance of players in listening trials. Experiment 1 has shown that players choose the matching item with a high accuracy with little practice, and this accuracy is believed to stay constant in this experiment.

Participants 25 participants were recruited through Amazon Mechanical Turk and compensated financially by receiving \$0.75.

Results

A linear mixed-effects model was conducted with the number of strokes as the dependent variable, and the interaction of the conditions with progressing gameplay as the fixed factor. Intercepts for subjects and items were used as random effects. We found a main effect of condition ($t = -3.31, p < 0.001$), showing that the number of strokes decreased by 1.06 ± 0.32 in the between condition. The interaction of between condition and the progressing gameplay proved to be significant ($t = -2.91, p = 0.0036$), lowering the number of strokes by -0.07 ± 0.02 , despite an increase of 0.07 ± 0.01 generally over time ($t = 5.2, p < 0.0001$).

We also performed a linear mixed-effects model to check for possible effects on comprehension of the interaction of conditions, progression of gameplay and the time it took the participants to make their choice. Random effects controlled

for participants, items and competitor images. The variable time was log-transformed to remove skew. The model revealed no effects on comprehension.

Discussion

As predicted, the number of strokes increased after roughly three speaking trials in the condition in which the competitor image was chosen randomly, but not in the condition in which comprehension only required discriminating between gender and age categories. (Fig. 6). Additionally, the number of strokes continuously dropped in the latter condition as the game progressed. This sudden change after a short acquisition time is in line with other studies, which have shown that novel strategies are picked up rapidly by the majority of the population after a short period of time (Steels, 2011). The number of strokes should stabilize in each condition if players continue playing. Figure 7 shows examples of successful squiggles with differing numbers of strokes and encoded features.

Both conditions performed equally well in the listening trial, which shows that both conditions adapted to their given environment successfully. Importantly, fewer details were needed in the between-category condition to achieve the same accuracy as squiggles with more complexity in the condition in which the competitor was chosen completely arbitrarily.

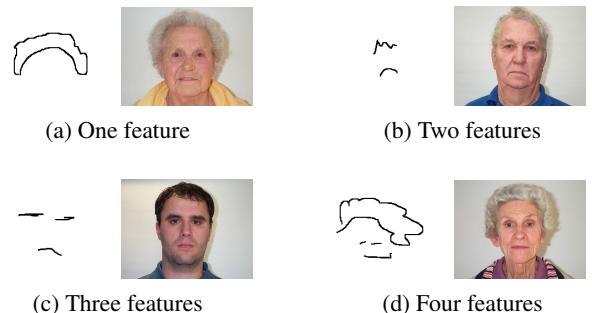


Figure 7: Different number of encoded features corresponding to the number of strokes a player drew.

General Discussion

Experiment 1 used a simple design in order to study the impact that visual and conceptual properties of have on the structure of squiggles drawn to communicate these objects. Input images differed in age and gender, and the produced squiggles varied in their structure according to these factors. Participants were able to adapt quickly to the given environment and performed well in comprehending squiggles from the start. The second experiment examined the impact of different comprehension conditions. An increase of listening trials at the start of the game and the introduction of different conditions led to varying squiggle structures across time and conditions. The referential environment had a profound impact on squiggle production. Specifically, we observed a decrease in squiggle complexity in the environment where the

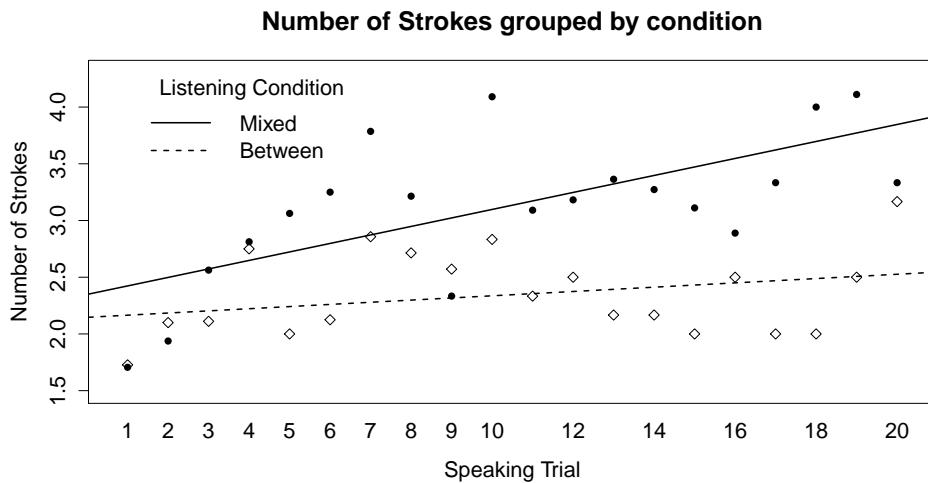


Figure 6: The evolution of the number of strokes over time across conditions

referents had to be distinguished categorically (although the participants were never informed of this constraint).

This method developed here to explore the evolution of communication in a simple graphical environment lends itself to numerous experimental manipulations. The impact of several parameters can be explored in this context. For example, a decreased time limit in the speaking trial will increase pressure on participants to draw more efficiently and to focus on fewer features. This may lead to a greater influence of the referential environment: Conditions in which categorical information (age and gender) is sufficient for successful communication lead to squiggles with simpler surface-structure, a likely starting place for the emergence of morphemes and compositional structure. The emergence of such morphemes could be sped up by using a system of immediate feedback, in which players are repeatedly informed how successful their squiggles perform in the listening trials, i.e. whether other players recognize their squiggles. Immediate feedback allows players to establish more common ground and has been shown to drastically increase communication performance (Garrod et al., 2007).

This study shows that humans rapidly create efficient communication systems in a novel shared environment by integrating contextual clues. Comprehension and production were tightly interconnected in this communication game, and the referential environment significantly affected the structure of the produced squiggles. This lends further support to the idea that communication systems adapt to the environment in which they are learned and used.

References

Chater, N., Reali, F., & Christiansen, M. H. (2009). Restrictions on biological adaptation in language evolution. *PNAS*, 106(4), 1015–1020.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, 31(5), 489–558.

Dale, R., & Lupyan, G. (2010). Squiggle: Large-Scale Social Emergence of Simple Symbols. In A. D. M. Smith, M. Schouwstra, B. de Boer, & K. Smith (Eds.), *The Evolution of Language: Proceedings of the 8th International Conference* (pp. 391–392). New Jersey: World Scientific.

Fay, N., Garrod, S., Roberts, L., & Swoboda, N. (2010). The interactive evolution of human communication systems. *Cognitive Science*, 34(3), 351–386.

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

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

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *PNAS*, 105(31), 10681–10686.

Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PloS one*, 5(1), 1–10.

Meir, I., Sandler, W., Padden, C., & Aronoff, M. (2010). Emerging sign languages. In M. Marschark & P. E. Spencer (Eds.), *Oxford Handbook of Deaf Studies, Language, and Education, Volume 2* (pp. 267–280). Oxford: Oxford University Press.

Minnear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*, 36(4), 630–633.

Steels, L. (2011). Modeling the cultural evolution of language. *Physics of Life Reviews*, 8(4), 339–356.

Wray, A., & Grace, G. W. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, 117(3), 543–578.