Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Differences in learnability of pantomime versus artificial sign: Iconicity, culturalevolution, and linguistic structure

Permalink

https://escholarship.org/uc/item/48r6q2td

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 41(0)

Authors

Delgado, Tania
Coulson, Seana

Publication Date

2019

Differences in learnability of pantomime versus artificial sign: Iconicity, cultural evolution, and linguistic structure

Tania Delgado (tadelgad@ucsd.edu)
Cognitive Science Department, 9500 Gilman Drive
La Jolla, CA 92093 USA

Seana Coulson (scoulson@ucsd.edu)
Cognitive Science Department, 9500 Gilman Drive
La Jolla, CA 92093 USA

Abstract

One of the central goals of language evolution research is to explain how systematic structure emerges. A cultural evolutionary approach proposes that the systematic structure of language arises from the use and transmission of language. Motamedi and colleagues (2016) investigated the influences of these forces on the evolution of language by generating an artificial sign language in the lab. Over several generations of new learners and their interactions, an initially unsystematic set of silent gestures developed markers for functional categories of person, location, object, and action. Here we describe results of two studies that compared the learnability of solo-produced pantomimes versus signals that had been transmitted and used by interlocutors. In these studies, participants saw an artificial sign and judged whether an English translation matched or mismatched the meaning of the sign. In an event-related potential (ERP) study, we found that mismatches elicited larger negativities in the ERP than matches. However, those effects were most reminiscent of the classic N400 response in the evolved signs. This study provides a clearer view on how the mechanisms that drive language evolution change language to adapt to a learner's brain.

Keywords: artificial language learning; gesture comprehension; iterated learning

Introduction

All languages demonstrate systematic structure. From the smallest units of sound, to words and phrases, the elements of language are not independent. These elements are part of a structured system that allows infinite expressive power through the reuse and recombination of those elements. Systematicity is a property found across the world's languages, but how does this systematic structure of language emerge?

One answer to this question appeals to the forces of cultural evolution. Languages, like species, change over time and are subject to similar evolutionary processes found in biological evolution, such as variation, selection, and inheritance. In this view, language is under selectional pressures from human cognitive biases and adapted to suit the human brain (Christiansen & Chater, 2008). The nature of linguistic structure would then be a product of the learning and processing constraints that derive from underlying neural mechanisms.

One avenue for investigating the emergence of linguistic structure is to examine natural languages in the early stages of linguistic development. Although most communities have long-established languages, emerging sign languages such as Nicaraguan Sign Language (NSL) provide us with the opportunity to observe how linguistic features arise in a new human communication system. In the 1970s, the Nicaraguan government established a school for deaf children. These children, who communicated with their families via idiosyncratic systems of home sign, were brought together and organically created a novel sign language (Kegl, 1994).

In the case of NSL, each incoming cohort to the school has shaped the language and furthered its development (Goldin-Meadow et al., 2014). One example of the emergence and development of grammatical structure in NSL can be found in the use of spatial modulation to mark semantic roles in sentences expressing events with both an agent and a patient. Senghas (2003) found that signers from the earliest generation did not use the direction of spatial modulation in their interpretation of such sentences, whereas signers from the next generation made systematic use of spatial location to determine who the patient of the event was. The properties and structure of NSL thus changed as a function of transmission to learners of the next generation, as well as its use between signers who had already acquired the rules of the grammar.

Recent laboratory studies of artificial languages likewise suggest that the cultural evolutionary mechanisms of transmission and interaction play pivotal roles in the emergence of language and its change over time (Kirby, Cornish & Smith, 2008; Kirby, Griffiths, & Smith, 2014;

Tamariz, Cornish, Roberts, & Kirby, 2012). Motamedi and colleagues (2016) investigate the impact of interaction and transmission on the evolution of language by generating an artificial sign language in the lab.

In their study, participants in an initial "seed" generation were asked to innovate gestures for concepts that vary across six themes and four functional dimensions. These concepts were selected to share salient semantic features across a thematic category (Figure 1). Signs from the initial seed generation demonstrate high iconicity, use a lot of space, require a lot effort, are redundant, and use similar salient features of the theme (e.g. handshape that represents scissors cutting). Moreover, the seed generation signs do not contain features to distinguish across functional categories within a theme.

	Functional Dimension			
Thematic Dimension	**Person**	**Location**	**Object**	**Action**
Food	chef	restaurant	frying pan	to cook
Church	priest	church	bible	to preach
Photography	photographer	dark room	camera	to take a photo
Concert	singer	concert hall	microphone	to sing
Hair	hairdresser	hair salon	scissors	to give a haircut
Police	police officer	prison	handcuffs	to make an arrest

Figure 1: Chart of the 24 concepts from Motamedi et al. (2016)

In an iterated language learning paradigm, new sets of participants came into the lab and were trained on the gestures produced by the seed generation, and played a communication game using those gestures. The signs produced by one of these participants in a dyad was then passed on to two new participants as the training set. The process was repeated for five generations in a transmission chain. This design was intended to create pressure for participants to develop a way to communicate the different dimensions of category structure. Concepts from within a thematic category were similar such that a pantomime of each might be difficult to distinguish across the functional categories.

Motamedi and colleagues (2016) show that under the pressures of communication and transmission, highly iconic and lengthy manual signals signs change to become more efficient and less iconic. After several generations of interaction, the authors also found the recycling of gestures within a theme. Most impressively, Motamedi and colleagues (2016) found the emergence and retention of functional markers that make it possible to distinguish between concepts within a theme. For example, in one dyad, signers pointed at themselves to indicate that the subsequent gesture depicted a person.

Despite their iconic origins, many of the functional markers are not transparent to new learners, and must be learned as arbitrary constructs. In one artificial sign system, for example, the marker for action involved the raising of the right hand with the palm facing out. The emergence of functional markers after several generations of learners in this study is used as a proxy of the emergence of systematicity in linguistic structure.

The Present Study

Here, we examine whether the communicative advantages of the final generation signs outweigh the benefit of the iconicity in the signs from the seed generation. Accordingly, we present videos of gestures from Motamedi et al. (2016) in a word learning task in which we compare participants' ability to learn the meanings of the iconic seed generation signs versus those of the more language-like final generation signs. We are interested in the processing and learning of language-like artificial signs, thus we applied methods typically used to study processing of natural languages.

In our study, participants viewed signals from the artificial sign language followed by English words that either match or mismatch the signal's meaning. We focus on two different ways in which the word presented can mismatch the meaning of the sign. A Thematic Mismatch is a violation of the thematic category, (e.g. present the sign for hairdresser, then display the word "chef" on the screen), whereas a Functional Mismatch is a violation of the functional category (e.g. present the sign for hairdresser, followed by the word "scissors").

In manipulating the generation that the sign comes from, we are able to see if there are differences between learning improvised pantomimes versus the signs evolved in the lab. We expect that identifying a mismatch in the thematic violation cases would not be difficult for either seed signs and evolved signs, as all signs displayed some degree of iconicity, and were readily distinguishable between thematic categories, (e.g. food versus photography). However, we expect that identifying a functional violation would be more difficult because signs within a theme share many iconic features associated with their thematic category, and may not provide features that would allow a learner to distinguish between the four potential meanings.

In Experiment 1, we measured response times and accuracies in a behavioural artificial language learning task. In Experiment 2, participants complete the same task as in Experiment 1, while we measure event-related potentials (ERP) time-locked to the onset of the English translation of the sign. We are particularly interested in the N400, ERP component known to index difficulty associated with meaning processing or retrieval from semantic memory, and is produced reliably across a range of stimuli (Kutas & Federmeier, 2011). Even within 14 hours of instruction, second language learners show larger N400 responses to pseudowords compared to real words that were semantically related or unrelated to primes, indicating that limited exposure is sufficient for new language learners to gain sensitivity to lexical status and word meaning (Mclaughlin, Osterhout, & Kim, 2004). ERP studies allow for real-time indexing of brain activity and provide multidimensional data about stages of processing. Thus, the N400 component is an appropriate dependent measure to more precisely examine the learning of an artificial language, in such a way that is comparable to studies investigating the processing of natural language.

Experiment 1

In Experiment 1, we taught participants signs from the Motamedi et al. (2016) in an explicit language learning paradigm. We used a within-subjects design in which each participant learned 12 signs from the seed generation and 12 signs from the final generation. In this behavioural experiment, we measured accuracy and reaction time in making judgements about whether the sign and word presented on the screen matched. We predict that accuracy will be greater for final generation signs after participants have learned the mappings, and reaction times will decrease as participants learn the system. We also expect lower accuracy rates and slower response times for Functional Mismatches.

Methods

Participants

We recruited 38 healthy undergraduates (15 M, 23 F). All gave informed consent and received course credit for participating. English was the primary language of all participants. One participant was excluded for not completing the experiment.

Materials and Procedure

Each trial began with a fixation cross for 500ms, followed by the video that varied from 2 - 7 seconds depending on signal length. A word then appeared until a key press was made, with feedback displayed on the screen for 500ms until the next fixation cross. We used two different stimulus lists varied across participants so that each concept was conveyed once with a seed gesture, and once with a final gesture. Participants watched videos of signs from either the seed generation or final generation. After each video was played, a word was displayed on the screen. The word either matched or did not match the meaning of the previously shown sign. When the word was displayed on the screen, participants pressed a key to indicate whether or not the word matched the sign. Participants received immediate feedback after every response they made. Feedback was given by the words "correct" or "incorrect" presented on the screen, and an accompanying tone. The experiment comprised 4 blocks of 48 trials each.

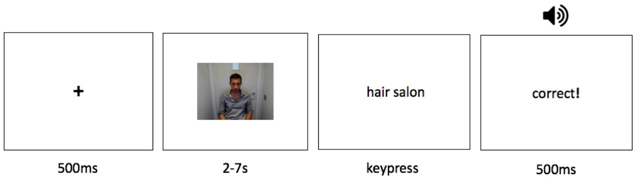

Figure 2: Example of a single trial.

Results and Discussion

Accuracy

A mixed effects logistic regression model was used to analyze the accuracy rate data. Models were constructed with the *lme4* package in R (R Core Team, 2013; Bates et al., 2015). Analysis involved construction of a generalized linear model to predict accuracy with experimental Block (First, Second, Third, Fourth), Generation (Seed, Final), and Condition (Match, Thematic Mismatch, Functional Mismatch) as categorical predictors, and all interactions. Models were fit with random intercepts for participants and for items (i.e. the videos that were played). Mean accuracy rates in each experimental category are shown in Figure 3. Model estimates are listed in Table 1. Analysis suggests accuracy rates improved as the blocks progressed. Experimental condition also impacted performance as accuracy rates were highest for Thematic Mismatches, lower for Functional Mismatches, with intermediate performance on the matches. The interactions between Condition and Block result because the learning curve was steeper for the more difficult Functional mismatches than the Thematic mismatches.

Participants' performance show that Functional Mismatches are more difficult to judge as being mismatches. The signs within a thematic category share many of the same features with respect to handshape and movement, such that differentiating between signs within a theme is ambiguous. Initially, participants perform worse in trials with final generation signs, which suggests that the markers contained in these signs are not transparent to new learners. There appears to be more arbitrariness to the form of a marker, i.e. an open hand facing palm forward denoting an action would not be considered an obvious association. However, after several trials participants quickly learn to map the marker to action verbs, as demonstrated by the increase in accuracy by the second block.

Table 1: Mixed effects logistic regression for accuracy rates.

	Estimates	t-value
Mismatch Type:		
Functional	-0.251	-8.94
Thematic	0.186	6.86
Generation	0.0134	0.600
Block	0.0520	9.10
Functional:Block	0.0554	5.51
Thematic:Block	0.0456	-4.51

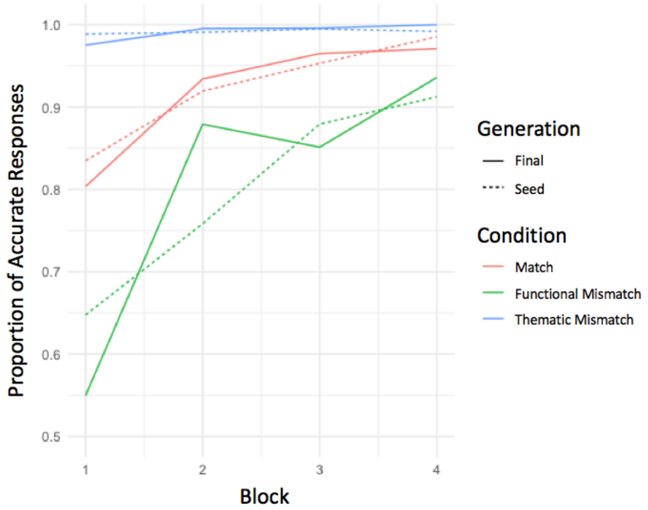

Figure 3: Accuracy rates between generation and condition across blocks.

Response Latency

To predict response latencies, we fit a linear mixed effects model in R (R Core Team, 2013) using the *lmer()* function of the *lme4* package (Bates et al., 2013). Predictor variables again included sign Generation (Seed, Final), Condition (Match, Thematic Mismatch, Functional Mismatch), and experimental Block (First, Second, Third, Fourth) and all interactions. Random intercepts were included for participants and item videos. Mean response latencies from each experimental category are shown in Figure 4 with model estimates listed in Table 2. Analysis revealed an interaction between Condition and Block, due to reaction times decreasing over the course of experimental blocks. Mismatch type also impacted performance as response latencies were consistently fastest for Thematic Mismatches, slower for Functional Mismatches, with intermediate performance for the Matches. Functional Mismatches also displayed the slowest average response latency across blocks, especially in the case of judging signals from the seed generation.

The results show that most of the learning of the mappings between sign and concept occurs during the first block of the experiment, as demonstrated by the slope of the response latencies from Block 1 to Block 2. As expected, participants respond faster to Thematic Mismatches since mismatches are easier to detect when the gestures produced clearly relate to different themes. Responding to seed signals is slower overall, which suggests that more processing occurs in deciding whether the signal matches the word presented. Seed signs are characterized as being longer in length, repetitive, pantomime-like, and lacking in defining features that would differentiate them from similar concepts. Between Blocks 3 and 4, there is a decrease in reaction time for decisions about final generation signals, suggesting that participants have mastered the meaning of the functional markers.

Table 2: Linear mixed effects model for response latency.

	Estimates	t-value
Condition	-0.573	-7.57
Generation	-0.573	0.726
Block	-0.208	-10.4
Condition:Block	0.132	4.91

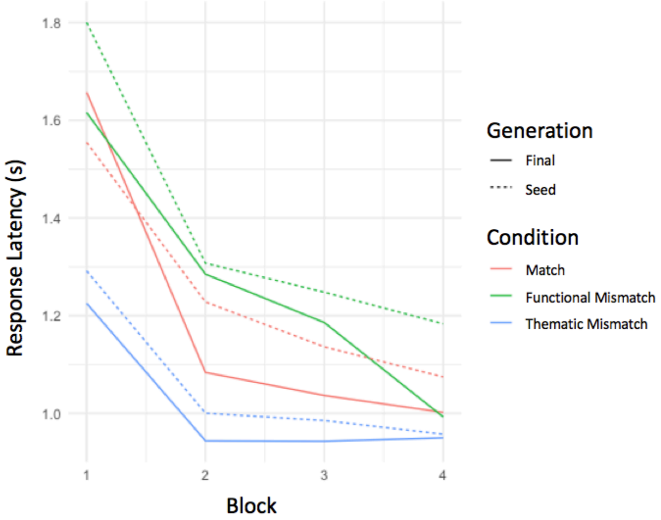

Figure 4: Mean response latencies between generation and mismatch type across experimental blocks.

Experiment 2

In Experiment 2, we measured neural responses in an artificial language learning paradigm. If the participant has learned the sign, we would expect mismatches to elicit a larger N400 response than matches. If final generation signs are indeed more learnable than those from the seed generation, we might expect to see larger amplitude N400 effects on words following the final generation signs than those following words from the seed generation.

Methods

Participants

We recruited 34 healthy undergraduates at UCSD (12 M, 22 F). All gave informed consent and received course credit for participating. English was the primary language of all participants. Two participants were excluded, one for excessive sneezing and sniffling, and one who was unable to complete the experiment within the allotted two hours. Participants completed surveys on handedness, neurological damage, and medication.

Materials and Procedure

Materials and procedure were adapted from the behavioural study outlined in Experiment 1.

EEG Data Collection

EEG was collected from 29 scalp sites using an ElectroCap mounted with electrodes. Scalp electrodes were referenced to the left mastoid. Blinks were monitored from an electrode below the right eye and referenced to the left mastoid. Horizontal eye movements were monitored via two electrodes placed beside each eye. Electrical impedance was reduced to less than 5 kohms. EEG was recorded and amplified using SA instrument bioelectric amplifier. The EEG was digitized at a sampling rate of 512 Hz. Recording took place in a dimly lit, sound-attenuated, electrically-shielded chamber. Participants were seated in front of a CRT monitor for stimulus presentation.

Results and Discussion

ERPs were time locked to the onset of potential meanings (viz. English words) presented after each signal. Mean amplitude was measured relative to a 100ms pre-stimulus baseline in two time windows: 300-500ms post-onset, intended to capture the N400 component, and 500-700ms post-onset, intended to capture the P600. In each interval, analysis involved repeated measures ANOVA with factors Condition (Match, Thematic Mismatch, Functional Mismatch), Generation (Seed, Final), Block (First, Second, Third, and Fourth), and two factors intended to capture the location of electrodes across the scalp, Hemisphere (Left, Right), and Region (Frontal, Frontocentral, Central, Centroparietal, Parietal, Occipital). Where relevant, the Greenhouse Geisser correction has been applied to p-values; however, for clarity, we report the original degrees of freedom.

Omnibus analyses revealed (among other effects) the presence of significant complex interactions with Block in both intervals (*N400:* Condition x Generation x Block x Hemisphere $F(6, 186) = 3.36, p < 0.05$; *P600:* Condition x Generation x Block $F(6, 186) = 2.76, p < 0.05$, Condition x Generation x Block x Hemisphere $F(6, 186) = 2.5, p < 0.05$), motivating separate follow-on analyses within each block.

N400 Analysis of ERPs in the first block revealed a main effect of Condition ($F(2, 62) = 8.2, p<0.05$), but no interaction with Generation ($F(2,62) = 1.03$, n.s.). By contrast, analysis of the second block suggested condition effects differed for signs from the seed versus the final generation (Condition, $F(2,62) = 18.4$ $p<0.001$; Generation, $F(1,31) = 4.2, p<0.01$; Condition x Generation, $F(2,62) = 3.26, p<0.05$; Condition x Generation x Hemisphere, $F(2, 62) = 7.8, p<0.01$). In the third block, Condition effects were present ($F(2,62) = 14.3, p<0.001$, but were similar for seed and final generation signals (Condition x Generation, $F(2,62) = 1.2$, n.s.). In the final block, Condition effects ($F(2,62) = 7.3, p<0.01$) displayed a different topographic profile following seed versus final generation signs (Condition x Generation x Region, $F(10,310) = 3.48, p<0.01$).

Figure 5 shows the topography of ERPs in the N400 interval for each type of mismatch following seed (upper panel) and final (lower panel) generation signs. Whereas the seed generation mismatches display a right frontal maximum reminiscent of ERPs to imageable words (see, e.g., Swaab, Baynes, & Knight, 2002), the topography of the final generation mismatches resembles the classic N400 that results from associative priming (e.g., Steinhauer, et al., 2017).

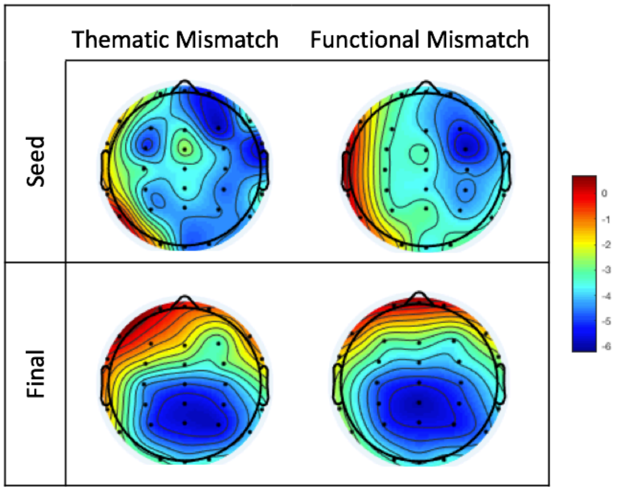

Figure 5: Difference in amplitude for latency between 300-500ms.

P600 Among other effects, follow up analyses revealed the presence of complex interactions between Condition, Generation, and topographic factors in blocks 1, 2, and 4 (*Block 1:* Condition x Generation, $F(2 ,62) = 4.55, p<0.01$; *Block 2:* Condition x Generation, $F(2,62) = 4.27, p<0.05$, Condition x Generation x Hemisphere, $F(2,62) = 7.23, p<0.001$; *Block 3:* Condition x Generation x Region $F(10,310) = 1.99$, n.s.; *Block 4:* Condition x Generation x Region, $F(10,310) = 6.96, p<0.001$). Figure 6 shows the topography of mismatch effects (match – mismatch) 500-700ms following seed and final generation gestures.

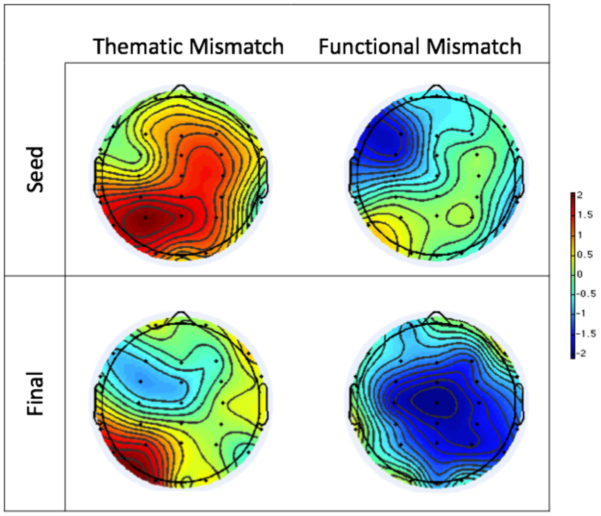

Figure 6: Difference in amplitude for each condition compared to the matches in the 500-700ms window

Figure 7 shows ERPs recorded at Pz, a parietal site where N400 and P600 are typically prominent. In the first half of the study, the N400 dominates the ERP response to these words, with more clear differentiation between the three conditions being evident in the seed generation signs. In the latter half of the study, N400 effects are overlapped by late positivities related to the task of classifying the word as a match or a mismatch. Following both seed and final generation signs, matches elicit a positivity that peaks earlier than the thematic mismatches (viz., seeing "chef" after the sign for hairdresser). For functional mismatches (viz., seeing "hair salon" after the sign for hairdresser), however, ERPs in the seed generation are more similar to the matches, whereas functional mismatches in the final generation are more similar to the thematic mismatches.

General Discussion

Here, we examined the ways in which a culturally-evolved artificial language may be advantageous to learn, in comparison to a system of individually iconic communication signals that lack internal systematicity. We found that signs from the more evolved system included both a consistent and concise iconic gesture to indicate thematic category, and a gesture that indicates whether the concept is a person, place, object, or action. Although the behavioral study suggested participants readily learned both the evolved final generation signs and the less systematic seed ones, the real-time brain response revealed processing differences for the two kinds of signs.

Our ERP study revealed a classic N400 response to signs from the final generation, indicative of semantic processing. By contrast, the iconic seed generation signs elicited concreteness effects that suggested participants exploited a learning strategy that involved mental imagery. Moreover, the brain response to final generation signs suggested participants could distinguish between closely related concepts such as hairdresser and scissors, whereas such concepts were treated identically in the seed generation.

Previous studies have also found that when used in a referential or communicative game, signs representing concepts from a set of shared semantic relations become more arbitrary, schematized, and systematized across dimensions (Theisen, Oberlander, & Kirby, 2010). We see that the introduction of a system of schematized signs influences how the meaning signs are retrieved from memory via ERPs to violations in signal-meaning pairings.

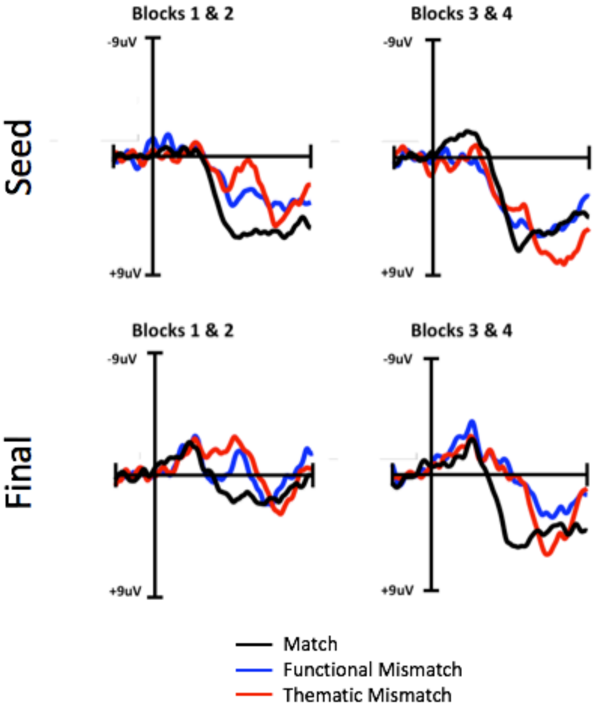

Figure 7: ERP waveforms recorded at electrode site Pz.

A recent study by Nölle and colleagues (2018) demonstrated how individuals use context and the environment to shape the signals they use together. The authors found that interlocutors were more likely to produce systematically-related signals rather than signals that refer to some idiosyncratic feature of the referent, even when both strategies were afforded by the environment. In the present study, we found that the brain's real time response displayed a greater sensitivity to subtle distinctions within a thematic domain for meanings conveyed by final generation signs that contained the functional markers. Systematic signs are easier to remember and rely more on abstraction to identify like features that can be referred to similarly.

Our current design adapts videos generated in a previous study as stimuli. This choice may introduce confounds relating to processing and working memory, as all seed generation signs were longer than the final generation signs. The seed signs are highly iconic pantomimes of actions associated with the theme, thereby resulting in longer signals that lack specificity. Consequently, the differences we found between learning seed and final signs might reflect differences in length of seed versus final generation signs, differences in the degree of structure, or some combination. Future research should seek to unconfound these factors.

Results of the present study support that artificial language shaped by interaction and transmission is more learnable. As such, it is in keeping with research that reports differences in the brain response in learners of another culturally-evolved artificial language (Verhoef, Walker, Marghetis, & Coulson, 2018). Signs evolved through interaction and transmission display systematic structure, and this systematic structure better suits the learner's brain.

References

Christiansen, M. H. & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, 31(5):489-509.

Goldin-Meadow, S., Brentari, D., Coppola, M., Horton, L. & Senghas, A. (2014). Watching language grow in the manual modality: Nominals, predicates, and handshapes. *Cognition*, 136, 381–395.

Kegl, J. (1994). The Nicaraguan sign language project: An overview. *Signpost*, *7*, 24-31.

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

Kirby, S., Griffiths, T., & Smith, K. (2014). Iterated learning and the evolution of language. *Current Opinion in Neurobiology*, *28C*, 108–114.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual review of psychology*, *62*, 621-647.

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature neuroscience*, *7*(7), 703.

Motamedi, Y., Schouwstra, M., Smith, K., & Kirby, S. (2016). Linguistic structure emerges in the cultural evolution of artificial sign languages. In *The Evolution of Language: Proceedings of the 11th International Conference. New Orleans: EVOLANG11* (pp. 21-24).

Nölle, J., Staib, M., Fusaroli, R., & Tylén, K. (2018). The emergence of systematicity: how environmental and communicative factors shape a novel communication system. *Cognition*, *181*, 93-104.

Senghas, A. (2003). Intergenerational influence and ontogenetic development in the emergence of spatial grammar in Nicaraguan Sign Language. *Cognitive Development*, *18*(4), 511-531.

Steinhauer, K., Royle, P., Drury, J. E., & Fromont, L. A. (2017). The priming of priming: Evidence that the N400 reflects context-dependent post-retrieval word integration in working memory. *Neuroscience letters*, *651*, 192-197.

Swaab, T. Y., Baynes, K., & Knight, R. T. (2002). Separable effects of priming and imageability on word processing: an ERP study. *Cognitive Brain Research*, *15*(1), 99-103.

Tamariz, M., Cornish, H., Roberts, S., & Kirby, S. (2012). How generation turnover and intelocutor negotiation affect language evolution. In *The Evolution of Language: Proceedings of the 9th International Conference on the Evolution of Language: World Scientific* (pp. 555-556).

Theisen, C. A., Oberlander, J., & Kirby, S. (2010). Systematicity and arbitrariness in novel communication systems. *Interaction Studies*, *11*(1), 14-32.

Verhoef, T., Walker, E., Marghetis, T., & Coulson, S. (2018). Neural measures of sensitivity to a culturally evolved space-time language: shared biases and conventionalization. In *Proceedings of the 40th Annual Meeting of the Cognitive Science Society*. Madison, WI: Cognitive Science Society (pp. 1151-1156).