

Linguistic Modality Affects the Creation of Structure and Iconicity in Signals

Hannah Little (hannah@ai.vub.ac.be), Kerem Eryilmaz (kerem@ai.vub.ac.be)
& Bart de Boer (bart@arti.vub.ac.be)

Artificial Intelligence Laboratory, Vrije Universiteit Brussel
Pleinlaan 2, 1050 Brussel, Belgium

Abstract

Different linguistic modalities (speech or sign) offer different levels at which signals can iconically represent the world. One hypothesis argues that this iconicity has an effect on how linguistic structure emerges. However, exactly how and why these effects might come about is in need of empirical investigation. In this contribution, we present a signal creation experiment in which both the signalling space and the meaning space are manipulated so that different levels and types of iconicity are available between the signals and meanings. Signals are produced using an infrared sensor that detects the hand position of participants to generate auditory feedback. We find evidence that iconicity may be maladaptive for the discrimination of created signals. Further, we implemented Hidden Markov Models to characterise the structure within signals, which was also used to inform a metric for iconicity.

Keywords: Linguistic structure; Combinatorial structure; Signal spaces; Iconicity; Hidden Markov Models

Introduction

One of the central interests to the field of language evolution is what initially motivated the emergence of structure in language, and how that structure manifests itself. Experimental work in laboratory settings using artificial languages is a dominant exploratory tool within the field, primarily focussing on signals built from pre-discretised blocks and the emergence of compositional structure, where meaningful elements combine to make bigger meaningful elements (e.g. Kirby, Cornish, and Smith (2008)). However, the scope of some experiments has started to shift to investigate the emergence of combinatorial structure, where meaningless building blocks combine to make meaningful units, using continuous signal spaces, for example Verhoef, Kirby, and De Boer (2014), which argues that phonemes emerged as the result of cognitive learning biases within cultural transmission.

In the current study, using a novel signalling space proxy, we manipulate both the structure of the signalling space, and the structure of the meaning space, to tease out how the dynamics between a signal space and a meaning space can motivate the emergence of structure and how that structure is defined. We are interested in how much structural emergence can be explained by physical mapping problems, in order to isolate what aspects of emerging structure are the result of more cognitive mechanisms.

Hypotheses

We are testing two related hypotheses: 1. The dimensionality of a signal space, or modality, will affect the emergence of signal structure, and 2. Ability to recourse to iconicity will also inhibit emergence of signal structure.

This first hypothesis is grounded in more than one observation. First, with more signal dimensions at our disposal

within a modality, the more signal distinctions that can be made. Hockett (1960) outlined that as soon as the amount of semantic distinctions outnumbers the number of signal distinctions, then we need a level of phonological (or combinatorial) structure. Second, the more similar the structure of a signal space to that of a meaning space, the easier it is to map meanings to signals directly, making iconic or compositional systems, where there is a one to one mapping between signal and meaning space, more likely than combinatorial ones, with meaningless building blocks. Dimensionality of a signal space will affect how similar it is to the structure of a meaning space. In natural languages, the sign modality has many more signal dimensions available than the spoken modality, and humans can visually perceive simultaneously presented aspects of a signal in a way that the auditory system can not (Sandler et al., 2011), meaning that it is easier to map signed signals onto highly complex meaning spaces than spoken ones. There are no known spoken languages without a level of combinatorial structure. However, recent evidence has shown that emerging sign languages can exist without a level of combinatorial structure, such as Al Sayyid Bedouin Sign Language (ABSL) (Sandler, Aronoff, Meir, & Padden, 2011). Considering the role of linguistic modalities with respect to their signal dimensionality in the emergence of structure is worth investigating.

The second hypothesis, that iconicity will inhibit emergence of signal structure, is related to the first, as when a signal space and a meaning space have matching dimensionality, then it is easier to iconically map one onto the other. De Boer and Verhoef (2012) built a mathematical model which explores how iconic signal-meaning mappings are optimal when signal and meaning spaces have matching dimensionality, and when there is a mismatch, then more conventionalised structural strategies are needed. However, it is important to keep in mind that the iconicity within this model was relative, where there is a correlation between signals and meanings in such a way that similar meanings will be represented by similar signals. Relative iconicity is usually not observable from individual signals without seeing the rest of the system. Examples include sound symbolism, where “glimmer” isn’t intuitively iconic, until one considers the correspondence between “gl” sounds (e.g. glitter, glam, glow) and shiny things. Relative iconicity is distinct from absolute iconicity, where a signal represents a referent directly, e.g. representing the form of a referent directly in a gesture (terminology from Monaghan, Shillcock, Christiansen, and Kirby (2014)).

Drawing from the evidence from ABSL, Sandler et al.

(2011) hypothesised that emerging languages using the manual modality may have more holistic signals than a spoken equivalent, before combinatorial structure becomes necessary, because of the ability to use “iconically motivated signs”. Further evidence on this matter comes from Roberts and Galantucci (2014) who present an experiment using a stylus on a surface that continually moved downwards in a constant motion so that participants could only manipulate the horizontal dimension (i.e. as in Galantucci (2005)). Importantly, Roberts and Galantucci (2014) manipulated the meaning space, rather than the signal space, in order to affect the iconicity of signals. Within the experiment, participants were asked to communicate a variety of meanings which were either squiggly lines, which could be represented through absolute iconicity with the modality provided, or circles coloured various shades of green, which could not be iconically represented. The experiment showed that the signals used for circles were made up from repeated elements, while the lines retained iconicity. However, this experiment only shows the effect of iconicity on structure at two extreme ends of the iconicity continuum, i.e. comparing absolute iconicity with an example where no mapping is possible.

Our Experiment

Our experiment tests if relative iconicity, rather than absolute iconicity, affects the emergence of structure within signals. Relative iconicity is much more prevalent in spoken language, where absolute iconicity is much more possible in the signed modality. In our experiment, we manipulate both the signal space that participants use to generate signals, and the meaning space which participants describe using their signals. In manipulating both, we affect the mapping between the two in different ways. We manipulate the dimensionality of both signal and meaning spaces, generating a dimensionality mismatch, which creates a mapping problem. Affecting the dimensionality of our signal space is a very simple way to model the differences between different linguistic modalities with different levels of dimensionality. Obviously, modalities for natural languages have a lot more dimensions than we use in this study, but as with any model, in order to isolate individual effects, simplicity is key. We also manipulate the meaning space, where meanings either differed continuously (making an intuitive mapping onto the continuous signal space), or the meanings were designed to be perceived as discrete.

Methods

Participants Participants were recruited at the VUB in Brussels. 55 participants took part in the experiment, 27 male, 28 female. Participants had an average age of 24. Participants were assigned to conditions randomly.

Signals Participants created signals using a “Leap Motion” device, an infrared sensor designed to detect hand position and motion. The sensor was used to detect a continuous signal space within which participants could gesticulate to produce audio signals. Depending on condition and phase within the experiment, signals could be manipulated in their pitch,

volume, or both. In phases where the signal could be altered in both pitch and volume, they were associated with the horizontal and vertical dimensions respectively. In phases where the signal could be altered in either pitch or volume, only one spatial dimension was used. Participants were given auditory feedback, but no visual feedback, other than seeing their hand position. Participants were given clear instructions on how to use the sensor, and time to get used to the mapping between their hand position and the auditory feedback in each phase of the experiment.

Conditions The conditions in our experiment differed in the meaning space which participants were asked to create signals for. The meaning space in both conditions consisted of a set of squares that differed along dimensions which were either continuous in one condition, or discrete in the other. In phases where the meaning space only differed on one dimension, five squares only differed in either size (in the continuous condition) or colour (in the discrete condition). In phases where it differed on two dimensions, nine squares differed either in both size and different shades of grey (in the continuous condition), or in both colour and texture (in the discrete condition) (see Fig. 1).

Procedure Participants were given instructions on how to generate signals using the leap motion device. They used one hand above the device to generate signals. There were three phases of the experiment, each phase consisted of a practice round and an experimental round, and each round consisted of a signal creation task and a signal recognition task.

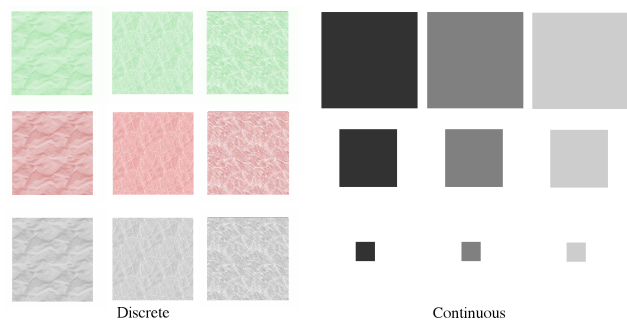


Figure 1: The meaning spaces used in phases 1:2 and 2:2 in the discrete condition and in the continuous condition.

Phases

In all phases participants saw the entire meaning space before beginning. In the signal creation task, they were presented with squares one by one and recorded signals using the leap motion. Participants were explicitly told which signal dimension(s) they were manipulating.

Phase 1:1 In the first phase participants were asked to create signals for a meaning space with 5 squares which only differed in one dimension (size or colour depending on condition). In this phase, participants could only manipulate the signal with one signal dimension, which was counterbalanced by randomly assigning participants to start with either pitch or volume. Which signal dimension the participant started

with was later controlled for in the analysis.

Phases 1:2 In phase 1:2, participants described a two-dimensional meaning space with the squares differing in size and shade in the continuous condition, or colour and texture in the discrete condition (Fig. 1). They used the same one-dimensional signal space used in phase 1:1 (see fig. 2).

Phases 2:2 In phase 2:2, participants were to describe the same two-dimensional meaning space as in phase 1:2, but this time with a two-dimensional signalling space where the signals differed in both pitch and volume (Fig. 2).

Counterbalancing Participants were randomly assigned to do either phase 1:2 or 2:2 first. Whether phases had matching dimensionality between signal and meaning spaces, or a mismatch, was used as a variable in our analysis. However, strategies will depend on what participants' have dealt with previously within the experiment. If they solve the dimensionality mismatch problem before being provided with the two-dimensional signal space, they may find it easier to continue with an already established strategy, than generate a new one taking advantage of both dimensions.

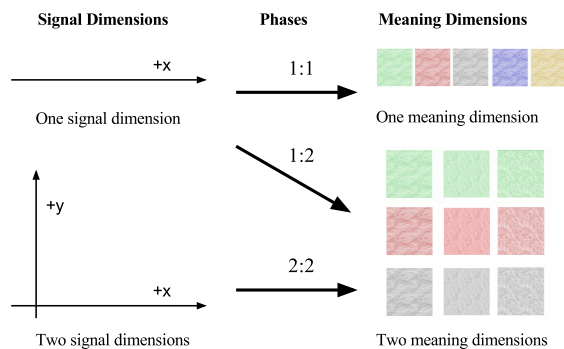


Figure 2: The mapping between signal space dimensionality and meaning space dimensionality in each phase using the meanings from the discrete condition (the continuous condition had the same mappings).

Signal Recognition task After each signal creation task, participants did a signal recognition task. They heard one of their signals, and were asked to identify its referent from an array of four randomly generated possibilities. They had immediate feedback about the correct answer. Their performance in this task was recorded for use in the analysis.

Post-experimental questionnaire After the experiment, each participant completed a questionnaire. Questions asked about what strategies were adopted in each phase of the experiment. The questions asked explicitly whether participants had strategies and what they were. Answers were free form.

Results

Post-experimental questionnaire When self-reporting their strategies, most participants gave similar answers within the continuous condition. Most choose to use pitch, volume or duration directly to encode size or shade using relative iconicity. For example, quiet to loud volumes used for light to dark shades respectively, or longer duration for bigger

squares. However, there was still some strategies involving different movement types, frequencies and speeds.

In the discrete condition, participants were more likely to attempt other forms of iconicity. Some associated the different colours with emotions or objects in the world like “a beating heart” or “the waves of the sea”, and then tried to make signals which corresponded to these things. However, participants were most likely to use patterns, speeds or frequencies of repeated elements.

From self-reporting, participants who saw phase 1:2 first were more likely to use the same signal dimension throughout than to change the strategy to take advantage of both dimensions. This was a significant association ($\chi^2(1) = 4.2, p < 0.05$). Also, participants were significantly more likely to perform better at the recognition tasks if they had strategies ($M=83\%$ correct), than if they didn't have strategies throughout ($M=52\%$ correct) ($t(19) = -5, p < 0.001$).

Signal Recognition Task Which condition participants were in had an effect on how well participants performed in the signal recognition task¹. Participants were significantly better at the recognition tasks if they were in the discrete condition ($M=82\%$ correct), than if they were in the continuous condition ($M=66\%$ correct) ($t(52.7) = -3, p < 0.005$).

The order in which participants received phases 1:2 and 2:2, and which signal dimension they started with (pitch or volume), did not reliably predict participants recognition of their signals. If a participant scored at chance level on the signal recognition task (1), they were disqualified from the rest of the analysis.

Measuring Structure

We started our analysis by generating standard deviations (SDs) for the coordinates of each signal trajectory in order to get some sense of how much movement there is within each signal, or whether more static strategies are used (which might be more indicative of relative iconicity). In the discrete condition, there was a tendency for SDs of signal trajectories to be bigger than in the continuous condition (23% on average), using a linear mixed effects analysis and controlling for participant and whether they started with pitch or volume as random effects, we found however that this effect was not significant ($\chi^2(1) = 1.9, p = 0.16$). However, we did find a significant effect of whether signals were produced in a phase with matching dimensionality (phases 1:1 and 2:2), or has mismatching dimensionality (phase 1:2), controlling for the same random effects ($\chi^2(1) = 8.6, p < 0.005$). Signal trajectories produced using the mismatch had a mean increase of 13.4% in their SD than those using mismatch.

We then created a measure for how predictable signal trajectories are. We quantised the signal coordinates using a k-means algorithm, in order to create a list of integer values representing a participant's entire repertoire of signals. With this, we estimated the marginal probability distribution of the

¹ If a participant scored at chance level on the signal recognition task (1), they were disqualified from the rest of the analysis.

points on each quantised trajectory. We then used these to calculate the conditional probabilities of individual points, and finally, the joint probability of whole signal trajectories by taking the negative logarithm of the product of first order conditional probabilities of the points on the trajectory. We found an effect of condition on how predictable signals were within a repertoire, using a linear mixed effects analysis and controlling for the same random effects as above ($\chi^2(1) = 11.3$, $p < 0.001$). The continuous condition showed more predictability than the discrete condition. Also, we found an effect of matching dimensionality, controlling for the same random effects ($\chi^2(1) = 5.8$, $p < 0.05$), signals produced using the mismatch had 16% more predictability than signals generated without a mismatch.

Hidden Markov Models

We used a Hidden Markov Model (HMM) with continuous Gaussian emissions to model the signal repertoires of participants. We treat HMM latent states as analogues for phonemes, and the emissions as analogues for the surface form. This allows us to use the number of latent states as an index of reuse (or structure) present in the repertoires.

For each phase of a run, we trained an HMM with all the signals used in that phase. Each signal is represented by an array of amplitude and frequency value couples that make up the signal. The number of latent states were determined by training multiple HMMs in parallel and picking the one with the lowest Bayesian Information Criterion (BIC) (see Algorithm 1).

Algorithm 1 HMM training and selection

```

1: function FITHMM(trajectories)
2:   hmm ← nil
3:   bic ← 999999
4:   nStates ← 2
5:   maxStates ← 30
6:   while nStates ≤ maxStates do
7:     for 1 : 100 do
8:       hmm' ← HMM(nStates)
9:       for trajectory in trajectories do
10:        hmm' ← BAUMWELCH(hmm', trajectory)
11:      if BIC(hmm') < bic then
12:        hmm ← hmm'
13:        bic ← BIC(hmm')
14:      nStates ← nStates + 1
return hmm

```

In order to validate that the model is relevant to participant success, we ran a mixed-effects regression on the participant signal recognition scores while controlling for condition, phase and participant. We used the normalized number of states (a real number in range [0,1], calculated by dividing the number of states of each HMM with the highest number of states in the group) as it was a slightly better predictor than the absolute number of states.

We found higher the number of states in an HMM, higher a participant's score ($R^2 = 0.604$, $\beta = 0.086$, $p < 0.01$). The

regression indicated significant random intercepts for participants and the interaction of condition (i.e. order of presentation) and phase, although adding random slopes did not improve the model.

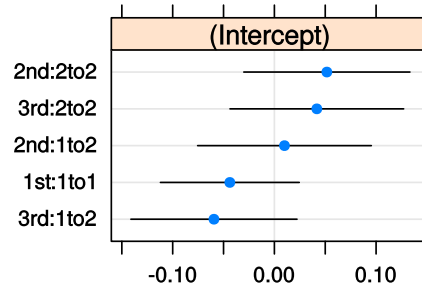


Figure 3: Random intercepts for each phase and condition pair. Conditions are represented as the phase presentation order).

The phase-by-phase analysis of the baseline number of states (as indicated by the random intercepts estimated for each phase) shows that the order in which the phases 1:2 and 2:2 are presented changes the expected number of states. If phases 1:1 and 2:2 are presented with an intermittent 1:2, there is a monotonous increase in the baseline number of states for each consecutive phase. However, when 1:1 and 2:2 are followed by 1:2, 1:2 ends up as the phase with the lowest baseline number of states in the experiment (Fig. 3).

Measuring Iconicity

In the continuous condition, it was easy to develop regression methods to demonstrate similarities between the signal space and the meaning space. Meaning dimensions were coded to reflect the continuous way in which they differed, e.g. the smallest square had a value of 1 while the biggest square had a value of 5. The signal trajectories were reduced to simple descriptive metrics, such as the mean coordinate value on each dimension, and the number of time frames reflecting its duration. Across all signals from all participants in the continuous condition, the mean coordinate value of the first dimension that a participant saw (either pitch or volume) was significantly correlated with shade of grey. We showed this using a mixed linear model, controlling for participant number and whether they started with pitch or volume as random effects ($\chi^2(1) = 341$, $p < 0.001$). Again, across all signals in the continuous condition, using a mixed linear model, controlling for the same random effects, duration of the signal was significantly correlated with the size of square ($\chi^2(1) = 103$, $p < 0.001$). Each square size increased the signal by 75 time frames ± 7 (std errors).

A problem arises when we try to use the same metrics to measure iconicity within the discrete condition, as it doesn't make sense to assign values to non-ordinal meaning dimensions. To tackle this, we developed a method that uses our HMM model in combination with the signal repertoire and the meaning space, to index iconicity. Some meanings in the

discrete condition are more similar to one another than others, e.g. a blue square is more similar to another blue square with a different texture, than it is to a green square with a different texture. We try to exploit this to generalise our notion of iconicity to the discrete cases, as well as making it algorithmic.

To measure iconicity within our results, we used Viterbi paths from the HMMs to reduce signals to a discrete sequence of states. This is an analogue of reducing a continuous speech signal of an uttered word to the string of phonemes which underlies it. A pair-wise distance matrix was then generated for the signal repertoire using Levenshtein distances between their Viterbi paths, representing how different each signal is in terms of number of “phoneme” changes necessary to get from one to the other. Then, the meanings were put into a pair-wise distance matrix with one another, to get a comparable measure of how many steps of “semantic” changes it takes to get from one meaning to the other. Mantel’s test of matrix correlation was then run between the two distance matrices to obtain an index of *how phonemic changes mirror semantic changes*, or how *relatively iconic* the repertoire is, in the form of a correlation coefficient between 0 and 1. If the null hypothesis that there is no correlation between the two matrices can be rejected we tagged that repertoire as iconic. Otherwise, we tagged it non-iconic, regardless of the estimated magnitude of the correlation.

Our preliminary analysis of this measure indicated we should expect it to produce some false negatives, i.e. iconic repertoires tagged as non-iconic², but our data is too limited to analyse this measure’s effectiveness adequately using classical statistics, so we built a Bayesian model to test it. The number of repertoires getting tagged “iconic” is represented by a Binomial distribution with a uniform prior on the parameter p_{iconic} , which is the p parameter for the distribution, or the probability of something getting tagged “iconic”. p_{iconic} was estimated separately for discrete (p_{iconic}^d) and continuous (p_{iconic}^c) cases, and for each phase.

The expected $p_{\text{iconic}}^c - p_{\text{iconic}}^d$ difference overall was positive, with 97.25% of the probability density above 0 (Figure 4). Although 0 does fall into the 95% credible interval, considering the overall distribution, it is reasonable to expect $p_{\text{iconic}}^c > p_{\text{iconic}}^d$. In other words, the continuous condition produces more relatively iconic inventories than the discrete case. Comparing phases within conditions using this measure did not indicate any significant trends.

Discussion and Conclusion

We were interested in two related hypotheses; 1. whether the dimensionality of a signal space or modality will affect the emergence of structure of signals, and 2. whether iconicity will inhibit emergence of signal structure.

First, we found support that dimensionality had an effect on the variance within signals, with signal trajectories produced

²Although it is beyond the scope of this paper, one reason for this is that our measure is confined to discovering linear correlations only.

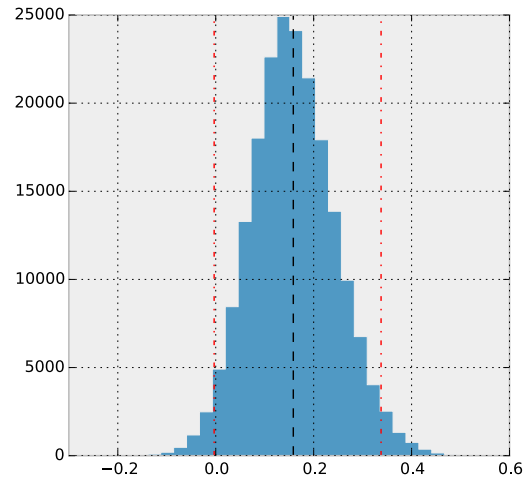


Figure 4: The posterior probability distribution of $p_{\text{iconic}}^c - p_{\text{iconic}}^d$ estimated by MCMC, as a histogram of the MCMC trace. The 95% credible interval is $[-0.006, 0.331]$, and the mean is 0.158. The figure shows that we can expect signals to be tagged iconic more often when they are formed in the continuous condition.

using the signal-meaning dimension mismatch having higher SDs. Greater variance indicates fewer, more distinct building blocks, which was also evident from the random slopes of the regression between number of HMM states and participant score. These results indicate not only that larger semantic spaces often cause more building blocks to be used, but also that the type of signals produced depends on the mapping between the semantic space and the signal space. Modulating the order in which matching and mismatching phases are presented changes the participant performance significantly, as shown in Figure 3. This effect arises from the strategy change required between phases with matching or mismatching spaces. When people gain experience in consecutive matching phases, the repertoire they bootstrap for the following mismatching phase becomes heavily compressed, as indicated by the low baseline for the number of states. However, when participants have to solve the mismatch problem first, there is an increase in the baseline with every phase, despite participants being able to employ their strategy from phase 1:2 in phase 2:2. This result contradicts what participants self reported in the questionnaire, where they have used the space in phase 2:2 maximally, no matter what the order of phases.

Second, we aimed to inform theories relating to the effect signal-meaning mappings have on the emergence of linguistic structure. We found support that when relative iconicity was possible, the majority of participants encoded size with duration, leaving them to encode shade with the signal dimension they were first exposed to. However, in order to compare the iconicity present in the continuous condition with that present in the discrete condition, we developed our own iconicity index, using HMMs. We found that signal repertoires in the continuous condition were more often tagged as iconic than

in the discrete condition. However, we did have problems including false negatives and the failure to confidently establish a difference between the phases in the continuous condition, demonstrating the need for further work. To complement the results on iconicity, we found that signals produced in the discrete condition, where relative iconicity was not possible, had greater variance than in the continuous condition, and also had significantly less predictability. We had initially thought that signals and repertoires with more repeated elements (or structure) would be more predictable, as they would be comprised from elements repeated throughout a repertoire, in the same way that phonotactic rules make natural languages more predictable than random strings of phonemes. However, our results suggest that a static signal will always be more predictable than one with movement, so perhaps predictability is not the best measure for structure here.

Further to the evidence pertaining directly to our hypotheses, we found that participants were better at recognising their signals in the discrete condition, than in the continuous condition. One might think that having a one to one mapping between signal and meaning would make a signalling system more intuitive, and perhaps easier to be productive with. However, the pressure against iconic systems in the discrete condition may have pushed participants to make more exaggerated differences between their signals within their chosen strategies. Further to this, signals that rely on relative iconicity are likely to be easier to confuse with each other, making them maladaptive for discrimination between signals. This fits in with findings from Monaghan, Mattock, and Walker (2012) where sound symbolism was found to be beneficial to category learning, but not beneficial for learning individual words.

Our experiment has shown that the physical aspects of different linguistic modalities, or signal space proxies, can affect the structure which can emerge. These effects are very important to consider before we can isolate the cognitive effects which experimental work in language evolution is trying to characterise (Verhoef et al., 2014). We have developed a new paradigm to address these questions, as well as new methods to measure structure within continuous signals. However, HMMs still present two limitations; 1. HMMs do not explicitly model time spent in each state, which some participants used as a strategy, and 2. Gaussian HMMs do not emit signals that are continuous in the signal space, which is a feature of the signal space proxy we use. We plan to address these issues by using explicit duration and autoregressive HMM flavours, which will allow more thorough comparison of the model and the modelled repertoire, since such HMMs can emit passable, continuous signals with explicit timing.

We have also considered the nature of the structure which we have seen emerging in our study. In previous experimental work, artificial languages have been shown to emerge to mirror the structure in a given meaning space (e.g. Kirby et al. (2008)), which would be considered compositional structure as each building block is meaningful. Having such

a structured meaning space in our experiment has meant that participants have generated signal structure which corresponds directly to the meaning space, something which our post-experimental questionnaires also confirmed. We plan to run further experiments where there is less internal structure within the meaning space in order to perhaps generate something more analogous to phonological structure. How iconicity affects the emergence of structure at both a combinatorial and compositional level is something we are very interested in pursuing, and we are currently planning future signal creation experiments with further manipulations to the signal and meaning space, as well as exploring these ideas within the context of communication and transmission. We also plan to further develop our metrics and models for use in analysing the results of our experiments, as well as helping inform parameters for new experiments.

Acknowledgments

We would like to thank the financial support of the European Research Council project, ABACUS (283435).

References

- De Boer, B., & Verhoef, T. (2012). Language dynamics in structured form and meaning spaces. *Advances in Complex Systems*, 15(3), 1150021-11150021-20.
- Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive science*, 29(5), 737-767.
- Hockett, C. F. (1960). The origin of speech. *Scientific American*, 203, 88-111.
- 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.
- Monaghan, P., Mattock, K., & Walker, P. (2012). The role of sound symbolism in language learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(5), 1152.
- Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014). How arbitrary is language. *Philosophical Transactions of the Royal Society B*.
- Roberts, G., & Galantucci, B. (2014). The effect of iconicity on the emergence of combinatorial structure: an experimental study. In A. Cartmill Erica, S. Roberts, H. Lyn, & H. Cornish (Eds.), *The evolution of language: Proceedings of the 10th international conference (evolangx)* (Vol. 10, pp. 503-505). World Scientific.
- Sandler, W., Aronoff, M., Meir, I., & Padden, C. (2011). The gradual emergence of phonological form in a new language. *Natural language & linguistic theory*, 29(2), 503-543.
- Verhoef, T., Kirby, S., & De Boer, B. (2014). Emergence of combinatorial structure and economy through iterated learning with continuous acoustic signals. *Journal of Phonetics*, 43, 57-68.