DIFFERING SIGNAL-MEANING DIMENSIONALITIES FACILITATES THE EMERGENCE OF STRUCTURE

HANNAH LITTLE, KEREM ERYILMAZ AND BART DE BOER

Artificial Intelligence Laboratory, Vrije Universiteit Brussel
Brussels, Belgium
hannah@ai.vub.ac.be, kerem@ai.vub.ac.be, bart@arti.vub.ac.be

Structure of language is not only constrained by cognitive processes, but also by physical aspects of the signalling modality. We test the assumptions surrounding the role which the physical aspects of the signal space will have on the emergence of structure in speech. Here, we use a signal creation task to test whether a signal space and a meaning space having similar dimensionalities will generate an iconic system with signal-meaning mapping and whether, when the topologies differ, the emergence of non-iconic structure is facilitated. In our experiments, signals are created using infrared sensors which use hand position to create audio signals. We find that people take advantage of iconic signal-meaning mappings where possible. Further, we use trajectory probabilities and measures of variance to show that when there is a dimensionality mismatch, more structural strategies are used.

1. Introduction

Artificial language experiments have started to use continuous signal-space proxies to investigate the emergence of conventions, patterns and categories within signals (e.g. Verhoef, Kirby, and De Boer (2014) and Galantucci (2005)). However, studies have primarily focused on structural emergence being the result of cognitive processes within cultural transmission (Verhoef et al., 2014) or communication (Roberts & Galantucci, 2012). Here, we investigate how the mapping between the signal space and the meaning space influences the structure which emerges within signals. How signal structure is affected by different linguistic modalities in real world languages is important when considering how and why linguistic structure emerged. Also, when considering emerging structure in experiments using different artificial signal space proxies, it is very important to understand the effects that those proxies have on structure before attributing emerging signal structure to purely cognitive processes.

The dimensionality of a signal space is the number of ways which meaningful distinctions can be made using that signal space, e.g. in speech, voicing or place of articulation. As such, the dimensionality of a signal space will effect a) how quickly semantic distinctions outnumber the number of signal distinctions, and b) how difficult it is to map signals onto complex semantic spaces. Previously, De Boer and Verhoef (2012) used a model to demonstrate that signal-meaning mappings are optimal when signal and meaning spaces share the same number of dimensions, and that when there is a mismatch between signal and meaning spaces, then more structural strategies are beneficial.

2. Experiments

We experimentally test how differences in the dimensionality of both the signal space, and the meaning space, will have on signal-structure. We are interested in what happens when there is a mapping possible between the dimensionality of the signal space and the meaning space, and what happens when there is a mismatch.

2.1. *Experiment 1*

2.1.1. *Methods*

Participants
25 participants, recruited at the Vrije Universiteit Brussel (VUB) in Brussels, took part in the experiment; 10 male and 15 female, with a mean age of 24 (SD = 4.6).

The signal space
An infrared sensor (*Leap Motion*) was used to create auditory signals generated from the hand positions of participants. Signals differed in pitch, volume, or both. Signals with two dimensions (pitch and volume), were created by moving one hand within a two dimensional space, i.e. moving a hand vertically would affect the volume, while a hand moving horizontally would manipulate the pitch. Both pitch and volume scales were non-linear. Signals could not contain gaps. In different phases, participants could either manipulate signals by moving their hand within a horizontal dimension, vertical dimension or both. Participants had time to get used to the mapping between their hand position and sound.

Procedure
There were three phases in the experiment (Figure 1), each phase had a practice round and an experimental round. Each round had a signal creation task and a signal recognition task. Practice rounds and experimental rounds were the same. Only the data from the experimental round was used in the analysis. Participants saw the entire meaning space before each phase. Signals were recorded for the squares one-by-one. Squares were presented in a random order. Participants could play back signals, and rerecord if they were not happy.

Phase 1:1 (matching) The meanings differed along one dimension (size, divided in 5 levels) and the signal space was also one dimensional (either pitch or volume).

Phase 2:2 (matching) Participants described the two-dimensional meaning space (differing in size and shade), with a two-dimensional signalling space (pitch and volume).

Phase 1:2 (mismatching) Participants created signals for a two-dimensional meaning space (differing in size and shade). However, the signal space had only one-dimension (the same as they used in phase 1:1).

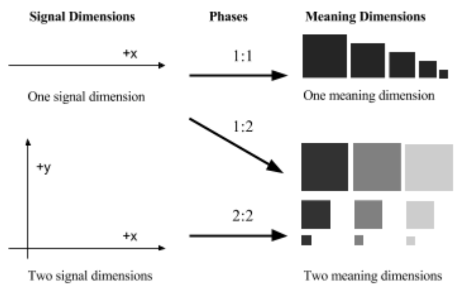

Figure 1. The phases used in the experiment.

Counterbalancing
Participants completed either phase 1:2 or 2:2 first.

Signal Recognition task
Participants heard a signal they had created, and were asked to identify its referent from an array of four randomly generated possibilities. They were given immediate feedback about the correct answer. This task worked as a pressure for expressivity, as participants knew they needed to recognise their own signals.

Post-experimental questionnaire
The questionnaire was free form and asked about the signal-creation strategies that the participant adopted during each phase of the experiment.

2.1.2. *Results*

Post-experimental questionnaire
Most self-reported strategies used pitch, volume or duration directly to encode size or colour where possible. Participants who saw phase 1:2 before phase 2:2 were more likely to self report using the same signalling strategy throughout, than to change the strategy to take advantage of both dimensions. This association was significant ($\chi^2(1) = 8.7, p < 0.01$).

Signal Recognition Task
Participants recognised a mean of 66% of signals correctly (25% expected by chance). When participants were incorrect, we were able to measure the distance between the answer they gave and the correct answer. Let m_{ij} define a meaning with size i and shade j. The distance between two meanings m_1 and m_2 is then the following:

$$D(m_1, m_2) = \sum_{k=i}^{j} |m_{1_k} - m_{2_k}|$$

Using this formula, we calculated the distance from the correct answer for both the actual data, and from data generated from choosing an answer at random. Comparing the actual data with the random data using a mixed effect linear model, and controlling for participant number as a random effect, and stimulus number as a fixed effect, we found that with incorrect choices produced in the matching phases, participants were closer to the correct square than if they had chosen at random ($\chi^2(1) = 5.5$, $p = 0.02$). However, in the mismatching phase there was no difference between actual incorrect choices and random incorrect choices ($\chi^2(1) = 0.01$, $p = 0.9$). Further, we found that the distance from the correct answer was much higher in the mismatching phases, than in the matching phases, indicating that participants were relying more on iconicity in the matching phases. We again tested this using a mixed effect linear model, and controlling for the same variables ($\chi^2(1) = 5.3$, $p < 0.05$).

Signal Creation Task

The data collected from the signal creation task consisted of coordinate values designating hand position at every time frame recorded. For this analysis, meaning dimensions were coded to reflect the continuous way in which they differed, e.g. the smallest square was 1 for size, and the biggest square was 5. Across all phases, the mean value of the first dimension that a participant saw in phase 1:1 (either pitch or volume) was predicted most strongly by shade. A mixed linear model, which included participant number as a random effect, and whether their starting dimension was pitch or volume as a fixed effect, showed this interaction to be significant ($\chi^2(1) = 341.4$, $p < 0.001$). The duration of the signal was predicted most strongly by the size of the square, with each step of size increasing the signal by 75.296 frames$\pm$7(std errors). The mixed linear model for this interaction, controlling for the same fixed and random effects, was also significant ($\chi^2(1) = 103.14$, $p < 0.001$). These correlations demonstrate a propensity for using iconic strategies. Size and duration are easy to map on to one another, and it makes sense that participants will more likely encode the remaining meaning dimension (shade) with the signal dimension they were first exposed to.

Standard deviations (SD) of the signal trajectories gave us a good idea of the amount of movement in a signal. Signal trajectories produced in the mismatch phase had higher SDs than signals produced in matching phases. Using a linear mixed effects analysis and controlling for participant number as a random effect, and whether they started with pitch or volume as a fixed effect, we found that this finding was significant ($\chi^2(1) = 4.5$, $p < 0.05$)

Probability of signal trajectories
Another measure we used was the predictability of each signal trajectory, derived from a participant's entire repertoire. We calculated the probability of each individual signal coordinate and used this to find the joint probability of each signal. We did this by taking the negative logarithm of the product of first order conditional probabilities of the coordinates on the signal trajectory. Using a mixed effects linear model and controlling for duration as a random effect, and size of square and participant number as fixed effects, we found that signals generated in phases with matched signal and meaning dimensionality were significantly more predictable than in phases were there was no match ($\chi^2(1) = 3.9$, $p < 0.05$). Signals produced in the matching phases had higher predictability.

2.2. *Experiment 2*

Experiment 2 tested the same hypothesis as experiment 1, but the design was altered to counter two problems with experiment 1, a) that duration was used as a dimension by some participants, meaning that if participants used duration, there wasn't a "mismatch" even with the 1:2 phase, and b) that participants produced signals for an entire meaning space in experiment 1 (5 or 9 meanings depending on the phase), meaning that generating holistic signals for each meaning was a possibility.

2.2.1. *Methods*

Participants
Participants were recruited at the VUB in Brussels. 25 participants took part in the experiment; 8 male and 17 female. Participants had an average age of 21.

Signals
As in the first experiment, there was a continuous signal space which used *Leap Motion*. In this experiment, signals could only be manipulated in pitch. Including duration, the number of signal "dimensions" could not be more than 2.

Meanings
The meaning space consisted of a set of squares which differed along continuous dimensions. In contrast to the first experiment, the number of possible squares from the meaning space outnumbered the number of squares presented to the participants, this created an incentive for participants to create more productive systems which extend to meanings they have not seen yet. Meaning bottlenecks have been shown to encourage the production of structure in experiments such as Kirby, Cornish, and Smith (2008).

Procedure

The procedure in experiment 2 was exactly the same as experiment 1 but with different phases.

Phase 1:1 Signals only differed in pitch. There were 6 squares which differed in 6 degrees of size. Squares were presented in a random order.

Phase 1:2 Participants were presented with 12 squares which differed along two dimensions, 6 degrees of size and 6 shades of grey stripes (See Figure 2.) This made a possible number of 36 squares which were chosen from at random.

Phase 1:3 Participants were presented with 12 squares which differed along three dimensions, 6 degrees of size, 6 shades of grey stripes and 6 shades of orange stripes (See Figure 2.) This made a possible number of 216 squares which were chosen from at random. Stripes were used as pilots showed that this made the squares more easily distinguishable than changing the shade of orange to reflect both darkness and redness.

Signal Recognition task
The signal recognition task was the same as in the first experiment.

Post-experimental questionnaire
Questions were about the strategies that participants adopted during each phase.

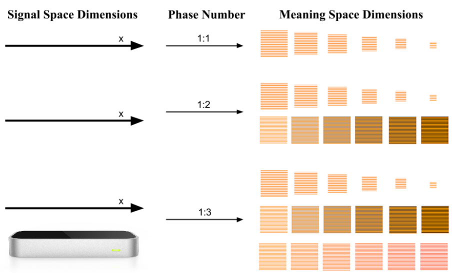

Figure 2. The dimensions used in experiment 2 in each of the 3 phases. All participants had phases presented in the same order: 1:1, 1:2, 1:3.

2.2.2. *Results*

Post-experimental questionnaire
In phase 1:1, participants self-reported encoding size with pitch or duration (80%). Participants tended to stick with the same strategy for size, but developed new strategies for new meaning dimensions. By phase 1:3, 56% of participants self-reported using a strategy which relied on movements or patterns.

Signal Recognition Task
Participants recognised their signals with a mean of 56% correct (again, chance level was 25%). When participants were incorrect, we were again able to measure the distance between their answer and the correct answer. Using a mixed effect linear model, and controlling for participant number as a random effect and square number as a fixed effect, we found that with incorrect choices produced across phases, participants were closer to the correct square than if they had chosen at random ($\chi^2(1) = 22.4$, $p < 0.001$) (see figure 3), the difference

between the actual and random result was significant within each phase. In later phases, there was more potential for incorrect distances being higher, because of the larger meaning space, meaning that a comparison between phases is not informative. However, the effect size for the comparison between the actual data and the random data in phase 1:3 was much smaller than in the other two phases, suggesting that in phase 1:3 there was less potential to rely on iconic strategies.

Signal Creation Task

The average duration of signals rose by about 20 frames each phase ($\chi^2(1) = 7.9$, $p < 0.005$)

As in experiment 1, meaning dimensions were coded to reflect the continuous way in which they differed. Across all phases, the size of square was the best predictor for the duration of the signal ($\chi^2(1) = 63.3$, $p < 0.001$). However, in this experiment, size was also the best predictor for the mean pitch of the signals ($\chi^2(1) = 15.7$, $p < 0.001$).

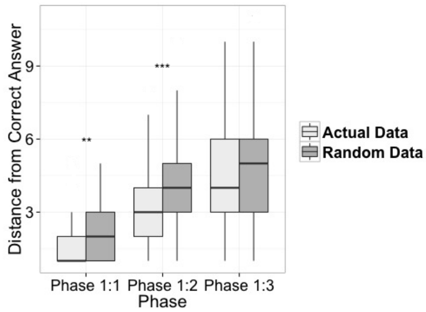

Figure 3. A graph showing the distance from the correct answer participants were in each phase when choosing incorrectly in the signal recognition task.

Looking at the SDs of individual signal trajectories, we see that the degree of mismatch affected the amount of movement in the signals. There was no significant difference between phases where there was no mismatch (Phases 1:1 and 1:2), in fact, the mean SD in these phases was nearly identical. However, the SDs from phase 1:3, where there was definitely a mismatch, were significantly higher than in the other two phases ($\chi^2(1) = 6.9$, $p < 0.01$).

Probability of signal trajectories

Calculating probabilities in the same way as in section 2.4m and using a mixed effects linear model, and controlling for duration and participant number as a random effect, and size of square as a fixed effect, we found that whether the signal was produced in a matching phase or not predicted how predictable a trajectory ($\chi^2(1) = 11.2$, $p < 0.001$). The log probability was closer to 0 (more predictable) in phase 1:1 (mean = 95), and got further away from 0 with each phase.

3. Discussion and Conclusion

In both experiments, we found correlation between structure in signal repertoires and structure in meaning spaces. This was particularly marked when signal and meaning spaces had the same number of dimensions. We also found more movement in signals in phases where there was a mismatch between signal and meaning spaces, suggesting more reliance on structural strategies.

Phases with matching dimensionalities produced signals which were more predictable given a participants entire repertoire, than signals produced within mismatching phases. This, again, may be indicative of the mismatching phases producing signals with more movement, as static, iconic, signals will be easier to predict.

We also found that in matching phases, when participants were incorrect, they were more likely to choose meanings which were closer to the correct meaning than if they'd chosen at random, again suggesting a reliance on iconic strategies.

We have shown that dimensionality, as a physical property of a signalling modality, will affect how, or ever if, structure will emerge. These findings are very important, both when exploring how structure emerges in signals produced using different linguistic modalities, but also when conducting experiments with signal-space proxies which have an effect on structure. Researchers must be careful to consider the physical effects of signal-space dimensionality when considering the emergence of structure. Importantly, understanding these effects will then allow us to isolate effects which are purely cognitive in nature.

Acknowledgments

Project supported by the ERC starting grant 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.

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.

Roberts, G., & Galantucci, B. (2012). The emergence of duality of patterning: Insights from the laboratory. *Language and cognition*, *4*(4), 297-318.

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.

© Little, Ery■lmaz, de Boer 2016
This work is licensed under a Creative Commons Attribution 4.0 International License.

©Little, Eryılmaz, de Boer 2016 - This work is licensed under a Creative Commons Attribution 4.0 International License.