UNIVERSITY OF AMSTERDAM

Efficient Coding in Speech Sounds: Cultural Evolution and the Emergence of Structure in Artificial Languages

T. Verhoef

Summary

Language is one of the most important features that separate us humans from the rest of the animal kingdom. This is why there is a great interest in discovering how language arose. This is not easy to find out, because there is not much tangible evidence to be found in this area. For a long time scientists could do little more than to use their imagination to develop theories about the evolution of language.

In the meantime, a lot has changed. Many researchers in the field of language evolution collaborate with researchers from other fields of study and new empirical methods have been developed. Geneticists for instance now search for unique genes that may explain human linguistic behaviour; computer modellers analyse and simulate evolutionary scenarios and interactions between individuals; linguists head into the field and study newly emerging (sign) languages; cognitive scientists and psychologists conduct experiments in which human participants learn or invent artificial languages and more.

The main question that is the focus of this thesis is: How did structure in speech arise? Speech is made up of basic building blocks: meaningless sounds are combined into words. Complex rules determine which combinations of sounds are correct in a language and which are not. How this property of language, *combinatorial structure*, emerged is still unclear. Some researchers assume the driving force behind it has to do with signal distinctiveness: the sounds used in language need to differ from each other maximally, otherwise words would sound too similar and we would be less well able to understand each other. Other researchers have proposed that principles of efficient coding play a role: a small set of sound primitives are reused and combined in a maximally efficient way. The results presented in this thesis show that the first assumption alone is not sufficient and that the second indeed seems to play an important role.

Experiments with humans and virtual robots

In my research I use two methods: experiments with human participants and computer simulations. The experiments can de compared to the game of ‘Chinese Whispers’ (or ‘Telephone’, ‘Broken Telephone’). In this

game a person whispers a message in someone else's ear and this message is passed on from person to person until it reaches the last person. The last person then says the message out loud and usually it is very different from the initial message, with often funny alterations. This game demonstrates in miniature what happens when languages are learned and reproduced repeatedly and transmitted from generation to generation. We call this process *cultural transmission* and this can be simulated in the lab by doing Chinese Whispers with entire (artificial) languages instead of single messages. Participants learn an artificial miniature language and are asked to reproduce this language. These reproductions are then passed on by asking the next participant to learn them. In this manner a chain of transmission is created and the transmitted language can be investigated. This method is called *iterated learning*. In the computer simulations, individual *agents* (virtual robots) interact with each other. These agents can produce and perceive sounds and learn an artificial language by dynamically updating their memory in response to interactions with others.

Figure 1: *Participant during whistle experiment in the studio.*

Evolving whistled languages

Natural language already has structure and regularities. How then, can we investigate the emergence of such structure in the laboratory with modern humans? In the experiments described in this thesis participants do not learn an existing spoken language, but a fictional miniature language. Participants cannot use their voice, but use an alternative, non-linguistic, device for sound production. In chapter 3 this device consisted of a two-dimensional interface in which scribbles on the screen were transformed into sounds. This appeared to be quite difficult to use and led to the use of whistles in subsequent studies. In chapter 4 and 6 the words of the artificial languages are whistled with a slide whistle. These whistled languages were transmitted and evolved in the laboratory with iterated learning. Figure 1 shows a participant with a whistle in the studio.

In chapter 4 the languages did not have meanings. Participants simply had to memorise a set of sounds. If we examine one of those evolved whistle languages, we can see that the set of sounds has gained a type of structure that is reminiscent of what we see in real languages. After some transmissions, a few basic building blocks can be identified and these elements are reused and combined in a systematic way. Figure 2 shows some of the sounds of such a language, where the pitch is plotted against time. Basic elements are clearly visible.

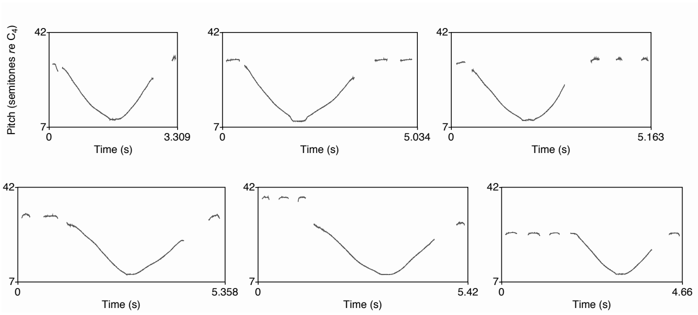

Figure 2: *Fragment of an evolved whistle language, plotted as pitch tracks. Basic elements, such as short level tones and falling-rising contours can be identified and they are systematically recombined.*

This is just one example, but the experiment was repeated four times. Each time a new chain of transmission was started with an unstructured source whistle language. Overall, the results show that the languages become easier to learn and more structured after a number of transmissions. This happens gradually and the participants are not aware of this.

In chapter 6 meanings were attached to the whistled signals. Participants were told that an alien space ship had crashed on earth and that the aliens needed help to repair their ship. To be able to help the friendly extraterrestrials, participants had to learn whistled words for alien space ship parts. Figure 3 shows a few examples of such 'space ship parts'. In this experiment with meanings, the results of the first whistle experiment were replicated. Even when the signals refer to meanings, and signal-meaning pairings could potentially be iconic and holistic, combinatorial structure emerged in all transmission chains.

UFO game experiments

As part of the analysis of the emerging miniature languages in the whistle experiments, another experiment was conducted which is described in chapter 5. The data for this experiment was collected both online and in Science Center NEMO and it consisted of a game in which participants had to save or destroy UFO's[1]. The UFO's contained aliens who spoke either the language of the good kind, or the language of the evil kind.

[1] This game was created by Jelle Zuidema and Vanessa Ferdinand

Figure 3: *Examples of 'space ship parts' used in the experiment.*

By listening to the sounds the aliens made, participants had to decide whether to shoot or save the UFO. Figure 4 shows a screenshot of the game.

The goal of this experiment was to investigate whether the evolved artificial languages from the previous experiments could be learned and distinguished by humans. The alien speech coming from the UFO's was constructed using the sounds from the whistle experiment. Two different conditions were created so that one group of participants got to listen to complete structures, while the others were exposed to random sounds, with no structure at all. If it is indeed true that the alien languages evolved in the experiments to become more learnable through an increase of structure, we would expect the first group, with exposure to the complete alien languages, to score much better at distinguishing between good and bad UFO's. This was indeed the case. On average, participants who could make use of the structure scored much higher than the other group.

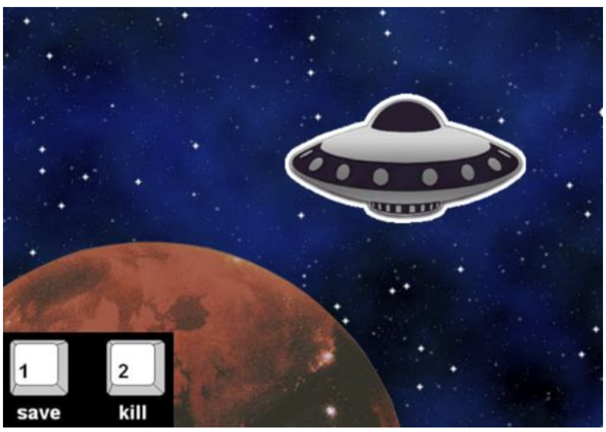

Figure 4: *Screenshot of the UFO game.*

Preservation of structure in populations of agents

Most of the chapters in this thesis focus on how linguistic structure emerges and develops when it is transmitted over generations. We have seen that the structures get simplified, more constrained and they become easier to learn. In addition to this focus on linguistic change, chapter 7 investigates how complexity and mutual intelligibility may be *preserved* over generations. The computer simulations described in that chapter involve experiments in which the emergence and development of artificial vowel systems is studied in populations of interacting agents. The main aim was to show how the preservation of complexity in vowel systems would be influenced when children learn faster than adults. Complexity appears to be preserved better over generations in populations where agents have such a *critical period*.

Conclusion

What this research can teach us is that structure in speech sounds can emerge as languages are repeatedly transmitted from generation to generation. Previously, the important influence of cultural transmission on structure in language has been studied for aspects of language such as syntax in detail, but for the study of complex compositional structure in phonology data was more limited. On the basis of those and other earlier findings, it has been pointed out that languages undergo their own evolutionary process and change gradually. Every generation of speakers changes the language a little bit without being aware of it. Within a language there is selection on structures that are learnable. Unnecessarily complicated rules or words will eventually disappear because speakers will not reproduce utterances they cannot learn. In this manner, cultural evolution causes languages to adapt to the human brain and become more learnable.

This is the first time iterated learning experiments have been conducted to study continuous sounds. My results provide additional support for the above mentioned ideas. Cultural evolution seems to be important in shaping not only compositional syntax, but phonological structure as well, because combinatorial structure in sounds can in principle emerge as the result of transmission and cognitive biases. The way in which structure emerges in the artificial languages in my experiments conforms with theories in evolutionary phonology that are based on principles of efficient coding. In general, iterated learning experiments appear to result in transmitted systems that become more compressible and more predictable. In future work theories on information-theoretic principles in neuroscience will be linked to these findings to gain a better understanding of the neurocognitive biases involved in the emergence of structure.