Complexity in speech: teasing apart culture and cognition

Bart de Boer

Vrije Universiteit Brussel

1. Introduction

Languages are more complex than strictly necessary for their communicative function. For example, the smallest repertoires of speech sounds found in language have about a dozen different contrasting phonemes (see table 1) whereas in the sample used for the UCLA Phonological Segment inventory Database (UPSID451, Maddieson, 1984; Maddieson & Precoda, 1990) the median number of speech sounds is 29 (with a first quartile of 23 and a third quartile of 36). This indicates that in general, languages tend

<table>
<thead>
<tr>
<th>phonemes</th>
<th>UPSID name</th>
<th>ISO 639-3</th>
<th>UPSID family</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Pirahã</td>
<td>myp</td>
<td>South American, Paezan</td>
</tr>
<tr>
<td>11</td>
<td>Rotokas</td>
<td>roo</td>
<td>East Papuan</td>
</tr>
<tr>
<td>13</td>
<td>Hawai’ian</td>
<td>haw</td>
<td>Polynesian</td>
</tr>
<tr>
<td>13</td>
<td>Nasioi</td>
<td>nas</td>
<td>East Papuan</td>
</tr>
<tr>
<td>14</td>
<td>Roro</td>
<td>rro</td>
<td>Malayo-Polynesian</td>
</tr>
<tr>
<td>14</td>
<td>Taoripi</td>
<td>tqo</td>
<td>Trans-New Guinea</td>
</tr>
<tr>
<td>15</td>
<td>Gadsup</td>
<td>gaj</td>
<td>Trans-New Guinea</td>
</tr>
<tr>
<td>15</td>
<td>Ekari</td>
<td>ekg</td>
<td>Trans-New Guinea</td>
</tr>
<tr>
<td>16</td>
<td>Ainu</td>
<td>ain</td>
<td>Isolate</td>
</tr>
<tr>
<td>16</td>
<td>Gugu-Yalanji</td>
<td>gvn</td>
<td>Pama-Nyungan</td>
</tr>
<tr>
<td>16</td>
<td>Bandjalang</td>
<td>bdy</td>
<td>Pama-Nyungan</td>
</tr>
<tr>
<td>16</td>
<td>Yidiny</td>
<td>yii</td>
<td>Pama-Nyungan</td>
</tr>
<tr>
<td>16</td>
<td>Dyirbal</td>
<td>dbl</td>
<td>Pama-Nyungan</td>
</tr>
<tr>
<td>16</td>
<td>Koiari</td>
<td>kbk</td>
<td>Trans-New Guinea</td>
</tr>
</tbody>
</table>
to use more than twice the number of phonemes that appears to be minimally necessary for full language. This makes it possible to define a notion of relative complexity: one language is more complex than another if a language user needs to learn more in order to use the one language than the other\(^1\). The question then arises why languages differ in relative complexity and how they become more complex.

This paper focuses on the second of these questions (as it is felt that it is more susceptible to empirical investigation) and more specifically on how one determines which aspects of linguistic complexity are due to cultural processes, and which aspects are due to cognitive biases. Cultural processes that shape language are the way language is used for communication and the way it is transferred from generation to generation or from group to group. Functional pressures, such as constraints on production and perception or information theoretical issues may play a role here as well. Some linguistic structures may be easier to understand or to produce and will therefore be preserved and transmitted better. Proponents of such a usage-based theory of language (Goldberg, 2003; Tomasello, 2003) explain linguistic structure from these cultural processes, while the underlying cognitive mechanisms are generally considered less important. In other words, even if the cognitive mechanisms would be rather different, still the same universals would be found.

Other linguists do explain properties of language from language-specific cognitive constraints. In their view, the way language is used plays a secondary role, and the structure of language is mainly due to properties of the brains of language users. An example of a model that explains structure of language from very language-specific cognitive biases is the principles and parameters model (Chomsky, 1981; Baker, 2002, 2003). In this paper, we will not argue for or against these positions, but

\(^1\) It should be noted that in this definition complexity of languages can only be compared for a given aspect of the language, such as the phoneme inventory size in the example.
rather address the issue of how one can investigate *empirically* what aspects of language may be due to cultural processes and what aspects of language may be due to cognitive biases. This in turn may help answer the question how languages become complex: is it the result of cultural and functional processes, or do we have cognitive biases that tend to make language more complex?

The paper addresses this issue from an evolutionary perspective. It adopts Dobzhansky’s (1973) point of view that “nothing in biology makes sense, except in the light of evolution”. What this means is that if one is interested in the *ultimate causes* of biological facts, one has to investigate how and why they have evolved. An evolutionary perspective has always been implicit in the study of language from a cognitive point of view: the properties of the human brain that linguists tend to focus on are the ones that are different from other animals, and that are considered essential for language. This is an implicit way of investigating which cognitive adaptations have evolved for language. Over the last twenty years or so, the evolutionary perspective on language has become increasingly explicit and influential (Pinker & Bloom, 1990; Hauser *et al.*, 2002; Fisher & Marcus, 2005; Fitch, 2005; Mufwene, 2008; Fitch, 2010).

The evolutionary perspective is doubly important for language: in order to understand potential biological adaptations for language, one has to study why and how they evolved biologically. In addition, if one wants to understand the properties of a particular language, one has to understand their history, i.e. how they evolved culturally. Cultural evolution is thus responsible for language change and usage-based emergence of structure, while biological evolution is responsible for the cognitive constraints that help humans learn language. The evolutionary perspective sees cultural processes and cognitive biases as two sides of the same coin, rather than as
mutually exclusive positions. However, when observing modern language, one only sees the final product of the interaction between culture and cognition. It is therefore difficult to tell which factor is responsible for observed linguistic phenomena.

This paper therefore introduces a recently emerging experimental paradigm – experimental cultural learning (Kirby et al., 2008; Galantucci, 2009; Scott-Phillips & Kirby, 2010). Experimental cultural learning attempts to re-create cultural evolution in a laboratory setting, and in this way to tease apart the effects of cultural processes and the effects of cognitive biases. It will be illustrated how this can be applied to the question whether complexity in phonemic systems is created by cognitive biases (feature economy) or rather by cultural processes (self-organization for distinctiveness).

In the next section, the recent debate about the role of biologically evolved cognitive constraints versus culturally evolved properties of human language is described in more detail. This puts the experiments that are described in the section after that in perspective. Finally we discuss how these experiments can be refined and how they may be applied to other aspects of language.

2. The controversy between biology and culture

Every linguist is aware that there is both a cultural and a biological component to language. Because there are so many different languages, and because children that are not exposed to a language do not spontaneously develop a language on their own\(^2\), it is clear that languages are products of culture. On the other hand, because only humans (and not, chimpanzees, elephants or dolphins, to name a few intelligent animals) learn language, it is clear that one needs a human brain to learn language. As

\(^2\)Interestingly, children not exposed to a language do appear to develop a new language if they grow up in a group, as illustrated by the case of Nicaraguan Sign Language (Polich, 2005).
in many areas of the study of human behaviour, there is a debate about the role of nature (biology) versus nurture (culture) in linguistics. However, because the above observations incontestably show that both culture and biology are important this debate is about the relative importance of both.

There are two independent questions in this debate: how specific to language are the cognitive mechanisms that are used in language and how much of the complexity and diversity of language is due to cultural factors versus biological factors. These questions are independent, because whether language cognition is specific to language or not does not determine whether cultural processes are most important in shaping language. Thus Chomsky (1957, 2007) proposes both highly language-specific cognition and an important role of these cognitive factors in shaping language. Tomasello (2003) on the other hand proposes highly general cognitive mechanism and an important role for culture. Baker (2003) proposes highly language-specific cognitive adaptations, but an important role for culture nevertheless. Griffiths and Kalish (2007) investigate a model that has a very general learning mechanism, and show that the properties of the language directly reflect the properties of the learning mechanism, thus giving less influence to cultural processes.

Even so, the position that there are highly language-specific cognitive adaptations tends to go together with the position that complexity in language is mostly determined by cognitive factors, while the position that language is based on general cognitive mechanisms often goes together with the position that complexity in language is mostly determined by cultural processes. This can perhaps be explained from the fact that general linguistics attempts to derive properties of cognition from observation of language(s) (as first widely advocated by Chomsky, 1957). It is then not extremely fruitful to look for evidence for language-specific cognition in language
if one does not believe that languages are determined by those cognitive properties. On the other hand, if one does not believe in language-specific cognition, then one is also less worried about the fact that properties of language may not be determined by cognitive biases. Although these combinations of points of view are therefore methodologically understandable, they are in fact independent, and which combination obtains in reality is an empirical question. Nonetheless, linguistics does seem to oscillate between these two positions.

Recently, the position that there are no language-specific cognitive adaptations and that languages are shaped by cultural processes appears to be gaining in popularity. Evans and Levinson (2009) have investigated many proposed universals of human languages and find exceptions for almost all the universals they investigate. They conclude that language universals are a “myth” and therefore that they cannot reflect strict innate restrictions of the human capacity for language. They do find that some patterns occur (much) more often than other possible patterns, but they explain this as the result of cultural evolution under functional pressure. In reaction to their article Tomasello (2009) has said “Universal grammar is, and has been for some time, a completely empty concept”.

To give further arguments against the possibility of evolved language-specific cognitive adaptations, Christiansen, Chater and Reali (Christiansen & Chater, 2008; Chater et al., 2009) have shown that non-functional biological adaptations for language cannot become fixed through biological evolution, because language changes too quickly through cultural evolution. Because cultural evolution works much faster than biological evolution, language is a moving target. Although in biology it is possible for acquired behaviour to influence the biological makeup of an organism through the Baldwin effect (Baldwin, 1896) this requires the environment in
which the behavior is acquired to be stable for a very long time. The Baldwin effect has been proposed by linguists as a way in which genetic adaptations to language can emerge (Pinker & Bloom, 1990) but Christiansen and Chater (2008) argue that language changes too quickly for this to work.

It must be noted that Christiansen, Chater and Reali (Christiansen & Chater, 2008; Chater et al., 2009) do allow for the possibility that properties of language that tend to recur because they are functionally useful may be stable enough that they influence the biological evolution of the brain. However, in the way their work is cited, for example by Evans and Levinson (2009) it often appears that this nuance is lost, and that it is impossible for language-specific cognitive adaptations to evolve at all. This leads to the impression that biological evolution has not played an important role in the evolution and complexification of language.

Interestingly, a very similar point of view is advocated by Chomsky (2007) when writing about evolution, even though (as pointed out above) Chomsky has advocated the importance of language-specific innate properties of the brain for understanding language. His point of view is that recursion is a crucial and language-specific property of human cognition (as pointed out in Hauser et al., 2002). However, because there is not such a thing as partial recursion (according to Chomsky, 2007), there is no evolutionary scenario by which recursion could have evolved through incremental improvements. As Chomsky (at least from his minimalist perspective) considers other aspects of language as peripheral, he therefore does not see an important role for biological evolution in the emergence of language, rather he proposes that language may be the result of an evolutionary accident or of exaptation from different functions (e. g. complex cognition). The unexpected convergence of viewpoints of researchers starting from very different perspectives about the lack of
importance of biological evolution for understanding language has been remarked upon by Berwick (2009). Of course, many linguists have a more nuanced position, and see language as the result of complex co-evolution between culture and biology (e.g. Mufwene 2013) and this is the position taken in this paper as well.

In fact, discarding biologically determined and language-specific cognitive mechanisms risks throwing the baby out with the bathwater. If there are no language-specific cognitive mechanisms, then why would one study language from a cognitive point of view at all? It would then be much more efficient to investigate the general cognitive mechanisms separately, and leave language to descriptive linguists. However, Chater et al. (2009) leave open the possibility that functionally relevant cognitive adaptations can evolve. In addition, it may be possible that linguistic cognition is less directly reflected in properties of language than Evans and Levinson (2009) appear to assume, or that there are properties of language that are universal, but that have been considered uninteresting by most linguists (e.g. de Boer 2014). Therefore, it may be too early to declare language-specific biological adaptations irrelevant.

One problem is that the search for language-specific cognitive adaptations appears to have focused on properties of language that do not have a clear function. This is understandable if one searches for language-specific cognitive adaptations: universals of language that cannot be explained from functional factors must be attributed to cognitive factors. However, this does not mean that properties of language that have a clear function, such as a large lexicon, are not associated with corresponding cognitive adaptations. On the contrary, Christiansen, Chater and Reali’s (Christiansen & Chater, 2008; Chater et al., 2009) work appears to indicate that these are the only aspects of language that are stable enough over time to
influence biological adaptation. And indeed, coming back to the example of large lexicons, it would seem plausible that humans have adaptations to learning large lexicons rapidly. However, the fact that we need to look for properties of language that are functional, makes it even more difficult to tell the effects of cultural and biological evolution apart.

Another problem is that the interaction between cultural evolution and biological evolution makes language one of the most complex phenomena in biology (Steels, 2000; Smith et al., 2003; Beckner et al., 2009). In fact, in order to understand language completely, one must take into account that language is influenced by phenomena on three time scales (Kirby & Hurford, 1997). The shortest time scale is that of individual language learning, which takes place on the time scale of a decade. Then there is the time scale of cultural evolution and language change, which takes place on a scale ranging from decades to millennia. Finally, the slowest time scale is that of biological evolution which takes place on time scales of the order of thousands to tens of thousands of years. These three processes are not independent, but influence each other, as illustrated in figure 1. This means that it can be hard to tell apart whether properties of language are due to functional constraints, cultural evolution or
biological adaptations. This is one of the root causes of the enormous disagreement about these issues among linguists.

The newly emerging experimental paradigm for studying cultural evolution in a laboratory setting may help to address these issues empirically. This may help re-interpret linguistic data, which by its very nature is always the result of a long process of cultural evolution under biological constraints. Even in the very few cases were emergence of a new language could be observed when it happened (Senghas et al., 2004; Polich, 2005; Sandler et al., 2005), the situation was insufficiently controlled to be able to draw general conclusions. The next section will introduce this new paradigm and illustrate it with a case study.

3. **Culture and cognition in speech**

In order to illustrate how effects of cultural processes and cognitive biases can be disentangled, an experiment on the emergence of structure in acoustic signals (Verhoef & de Boer, 2011; Verhoef et al., 2011a; Verhoef et al., 2011b) and its theoretical background will be presented in some detail.

3.1. **Cultural and cognitive explanations of speech sounds**

Human language is characterized by the use of a limited number of building blocks in order to produce an unlimited number of utterances. Interestingly, the inventories of building blocks that are used are not random. The building blocks appear to be constructed using a set of basic distinctive features (Jakobson & Halle, 1956). For example, the consonants [b] and [p] differ in the feature “voicing”: [b] is voiced and [p] is voiceless, but they are the same for other features, for example they are both “labial” consonants. The vowels [i] and [u] differ in the features “backness” and
“rounding”, where \([i]\) is front and unrounded, while \([u]\) is back and rounded. They are the same however for the feature “height”, being both high vowels. Some building blocks and some feature distinctions are very common in human languages, others are very rare. Rare features and speech sounds are called “marked” by linguists. Ever since the conception of the idea of distinctive features, there has been a debate about whether this is due to cognitive constraints or whether this is due to functional factors.

Jakobson and Halle (Jakobson & Halle, 1956) and Chomsky and Halle (Chomsky & Halle, 1968) among others have argued for the position that distinctive features reflect cognitive biases. They use observations from language acquisition and aphasia to argue that there are learning biases that favour certain speech sounds and features over others. Other researchers argue that there are functional reasons why sound systems are the way they are. Liljencrants and Lindblom (1972) show that optimization of distinctiveness leads to realistic vowel systems for small vowel system sizes. Stevens (1972) argues that certain speech sounds are inherently more robust and less subject to small errors in articulation than others, and are therefore preferred in language. More recent work using agent-based computer simulation has shown that optimization for distinctiveness in vowel systems can emerge through self-organization in a population (Berrah et al., 1996; de Boer, 2000). Apparently, innate distinctive features are not necessary to explain small inventories of speech sounds. However, because the models work less well for larger inventories, other factors must also play a role.

One factor appears to be feature economy (Clements, 2003, who also gives a historical overview of the concept) or the maximal use of available distinctive features (Ohala, 1980). If one analyses the sound systems of many languages, one finds that the features that distinguish speech sounds are generally not used to
distinguish just two speech sounds in the repertoire, but whole series of sounds. For example when nasalization is used in vowel systems, it is generally not used for only one vowel, but for a series of vowels. Similarly, consonants tend to occur in series: when voiced and voiceless plosives are used, generally these occur for all places of articulation used in the plosive system. Examples of the vowel system of French and the plosive system of Hindi are given in figure 2. Thus distinctive features tend to be used in a maximally economic way to create large systems of speech sounds.

It is possible that feature economy is the result of cultural processes, but it is equally possible that it is the result of a cognitive bias. Lending support to the interpretation of feature economy as the result of self-organization are the computer simulation results of Berrah and Laboissière (1999). These show emergent economic use of features in a population of simulated agents that have to communicate as effectively as possible. There is no explicit mechanism for economic use of features, and the authors therefore conclude that it is an emergent effect of functional constraints and cultural evolution. In a very different line of work however, Maye et al. (2008) show that infants appear to generalize features from familiar to unfamiliar contrasts. This appears to indicate that infants do have cognitive mechanisms for extracting distinctive features from speech data and for generalizing these features across different speech sounds. Because Maye et al. did not investigate feature
economy, one should be careful with extending their results to this area. Nevertheless, the mechanisms they uncover are an important prerequisite for an explanation of feature economy based on cognitive factors. There is therefore evidence that both cultural and cognitive processes may play a role in the explanation of feature economy. The challenge is therefore to experimentally tease apart the effect of culture and cognition.

3.2. **Experimental cultural learning**

The experimental paradigm of iterated (cultural) learning that has been developed over the last few years (Galantucci, 2005; Griffiths *et al*., 2008; Kirby *et al*., 2008; Smith & Kirby, 2008; Scott-Phillips & Kirby, 2010) may help to unravel the relation between culture and cognitive biases in language. Iterated learning is the repeated learning of culturally transmitted information in a group of agents, and usually focuses on human learning of language. The idea of studying the effect of repeated learning in a group was originally developed by computer modellers of the evolution of language. Two variants were proposed: Steels (1995, 1997, 1998, 2011) initially focused on interactions within a group of agents that in this way negotiate a shared communication system. This kind of cultural learning is alternatively called a language game, social coordination or horizontal transmission. Kirby and Hurford (Hurford, 1989; Kirby & Hurford, 1997; Kirby, 1999) initially focused on transfer of linguistic information across generations, which usually consisted of one agent. In their model a parent agent produces linguistic utterances that are learned by a child agent. After a while the parent is removed, the child becomes the new parent, and a new child is added to the population. This approach is alternatively called\(^3\) iterated learning.

---

\(^3\) As experimental iterated learning is a newly emerging paradigm, there is confusion about terminology. Kirby *et al* prefer to call all these approaches *iterated learning* and distinguish between *diffusion chains* (studying vertical transmission) and *social coordination* (studying horizontal transmission).
learning, vertical transfer, or a diffusion chain. Both types are illustrated in figure 3. After initial successes with these computer models, an increasing desire to link the simulations to real-world phenomena led to recreating them in a laboratory setting with human participants (Galantucci, 2005).

In such experimental iterated learning tasks, participants have to learn sets of signals produced by other participants in the experiment. The fact that stimuli are produced by other participants in the experiments instead of coming from a carefully set crafted by the experimenter is an important distinction between experimental iterated learning and more classical artificial language learning experiments. In most experimental iterated learning experiments, participants are required not just to learn and (passively) recognize a set of signals, but also to reproduce these signals.

Many variants are possible: experiments can either implement horizontal or vertical transmission (cf. Garrod et al., 2010 for a critical comparison). Signals can be

---

transmission). Other authors equate iterated learning with vertical transmission and call the general approach cultural learning.
used communicatively or not and they can have meaning or not. Experiments can be about strings of discrete symbols (letters) and thus address questions related to syntax or about continuously variable signals and thus address questions about phonology and phonotactics. These and other variants are reviewed in several recent papers (Kirby et al., 2008; Smith et al., 2008; Galantucci, 2009; Scott-Phillips & Kirby, 2010).

The results from iterated learning experiments can be analysed in two ways: by looking at either the system of signals that emerges from the experiments or by looking at the ways in which participants solve the problem of learning and reproducing sets of signals. Analysing the system of signals that emerges is comparable to analysing existing real languages: they are the product of cultural evolution, language-specific cognitive processes, and general cognitive processes combined. The ways in which participants learn and generalize utterances, however, make it possible to directly observe the individual behaviours that shape cultural processes. By manipulating the experimental setup, or by complementing iterated learning experiments with specially crafted artificial language learning experiments, it is possible to zero in on whether these behaviours are driven by functional factors or by cognitive factors and whether they are driven by language-specific or domain general cognition. These methods will be illustrated with a small case study below.

3.3. Iterated learning of speech

The work presented here is from Verhoef and de Boer (2011) and Verhoef et al. (2011a, 2011b, 2014). These iterated learning experiments investigate the emergence of combinatorial structure (the ability to combine a limited set of meaningless building blocks into an unlimited set of meaningful utterances) in systems of acoustic signals. Combinatorial structure is the ability to construct an unlimited number of
utterances from a limited set of speech sounds. In human languages this happens according to a set of learned rules that are generally referred to as the language’s phonotactic structure. As explained above, some researchers have claimed that combinatorial structure is the result of cultural processes, while others have proposed that it is due to cognitive biases related to language learning and language use.

The hypothesis that structure is (mainly) due to cultural processes makes radically different predictions from the hypothesis that structure is (mainly) due to cognitive processes. A mainly cultural process, as modelled for example in de Boer (2000) and Zuidema and de Boer (2009) results in slow changes that spread in the language community because the systems containing the changed sounds are better at avoiding confusion. A mainly cognitive process, such as feature economy discussed by Clements (2003) and maximal use of available distinctive features discussed by Ohala (1980), results in rapid emergence of (sets of) new building blocks and re-use of these building blocks in the utterances that make up the language. Verhoef et al.’s (Verhoef & de Boer, 2011; Verhoef et al., 2011a; Verhoef et al., 2011b; Verhoef et al. 2014) experiments directly investigate this issue.

These experiments investigate vertical transmission of small systems of initially unstructured signals. Participants are exposed to a set of 12 different signals and are then asked to reproduce them four times. Their last set of reproductions is used as training data for the next participant in the chain. If one of the signals that the participants produce is too close to one they have already made, this is flagged, and they are asked to retry to produce a signal. It should be noted that no meaning is associated with the signals. In order to avoid interference from pre-existing linguistic knowledge as much as possible, the signals are not spoken sounds, but are short whistles produced by means of a slide whistle. Because most participants are not
familiar with the use of a slide whistle, they can practice for a short period before the actual experiment. Four different chains of ten participants each have been investigated so far.

It turns out that systems of signals become significantly more learnable over the generations (as measured by Page’s trend test on the recall error of the participants, Verhoef & de Boer, 2011). Also, the signals in the systems become more similar (as measured by Page’s trend test on the average distance between a signal and its nearest neighbour in the set, Verhoef & de Boer, 2011). In addition, it appears that the entropy of the building blocks as used in the set of signals also decreases over the generations (Verhoef et al., to appear 2012).

The first observation would be expected for both the cultural and cognitive hypotheses: more distinctiveness predicted by the cultural hypothesis would make it easier for participants to keep signals apart, whereas more structured sets of signals predicted by the cognitive hypothesis would be more easily learned. The second observation is in contradiction with the cultural hypothesis, which predicts emergence of more distinctive signals, and therefore a larger distance between signals. The cognitive hypothesis does not really make a prediction here, although increased re-use of building blocks could make signals constructed out of these building blocks more similar. The third observation is predicted by the cognitive hypothesis. However, it has been shown that cultural processes can also result in sets of signals that appear to reuse building blocks (de Boer & Zuidema, 2010). Taken together, the observations do appear to indicate a problem for the cultural hypothesis, but it is necessary to look more closely at what happens during learning and reproduction in order to find out which hypothesis (if any) can be rejected.
When we look at individual reproductions of learned stimuli, it is much clearer that this involves a process of re-use and generalization of building blocks. Participants have a hard time learning the initial, unstructured set. However, when reproducing signals, they are forced to produce twelve distinct signals. This process results in subconscious generalization and re-use of observed patterns. Three processes by which participants do this are illustrated in figure 4. The first process is the combination of (parts of) one signal into a new signal. Verhoef and de Boer call this “borrowing”. The second process is reversing the pitch pattern of a signal and including this as a new signal in the new set of signals. Verhoef and de Boer call this “mirroring”. The third process creates new signals by repeating elements from existing signals. Verhoef and de Boer call this “duplication”. These processes are clearly indicative of (conscious or subconscious) manipulation of elements of learned signals and must therefore be based on cognitive processes. They are very different from the small and random shifts that are used in models of cultural processes, such as the simulations by Zuidema and de Boer (Zuidema & de Boer, 2009; de Boer & Zuidema, 2010). It must therefore be concluded that in this experiment, cognitive processes play a leading role in the emergence of structure.

Figure 4: Three processes through which experimental participants create new signals in their imperfect reproductions of a learned set of signals. In the left panel, borrowing is illustrated through the recombination of one signal of the learned set with the second half of another signal from the learned set. In mirroring, a signal of the learned set appears both in its original form and in a temporally mirrored form. In duplication, a signal from the learned set is analyzed as consisting of two elements. One of these elements is duplicated in reproduction. Signals were redrawn from (Verhoef & de Boer, 2011).
This does not mean that cultural processes do not play a role whatsoever. In fact, structure emerges only gradually over the experimental generations, indicating that structure is amplified by cultural transmission of the system of signals. In addition, the different chains of transmission from participant to participant appear to converge on differently structured sets of signals. It is currently tested experimentally whether human listeners can identify this culturally transmitted structure (Verhoef, 2013 ch. 5). In these experiments, participants (different from the ones who participated in the original experiment) are exposed to signals that emerged from a particular chain. Then they are tested with a mixture of signals that emerged in the same chain they were trained with (but that they have not heard before) and signals that emerged from different chains. It is measured whether participants can reliably distinguish between signals that emerged from their training chain and those that did not. These experiments appear to indicate that chains of culturally transmitted signals converge to distinguishable sets of signals (thus stressing that cultural effects are important in addition to the cognitive processes described above).

4. **Discussion**

This paper has attempted to illustrate a possible method, with origins in the study of language evolution, of teasing apart the role of cultural and cognitive processes in the emergence of complexity in language. It is a central question in linguistics which aspects of language are due to cultural processes, which aspects are due to language-specific cognitive processes and which aspects are due to general cognitive processes. In the case of general cognitive mechanisms, the evolutionary pressures that shaped these mechanisms have not been specific to language, whereas in the case of language-specific cognitive mechanisms, there must have been evolutionary pressures
specific to language. It is difficult to assess the role of these different mechanisms, as existing languages are the product of continuous interaction between the three mechanisms.

The paradigm of experimental cultural (or iterated) learning was developed by researchers of the evolution of language in order to study the effects of cultural processes in a laboratory setting. It consists of repeated interactions between participants in which language-like stimuli are learned, reproduced and (possibly) invented. In this way it can be studied directly how language-like systems change due to interactions between users of the system. Human behavior can then be compared with predictions both of models that are based on cultural processes and of those that are based on cognitive processes.

This approach has been illustrated for the case of emergence of complex combinatorial structure in acoustic signals. It was shown that human behaviour corresponds more closely to behaviour predicted by models based on cognitive processes (comparable to feature economy or maximal use of available distinctive features) than by models based on purely cultural processes (self-organization under pressure of distinctiveness). However, as the different learning chains converge to sets of signals that are structured according to different rules, cultural processes must play an important role as well. If only cognitive processes would play a role, then one would expect the learning chains to always converge to very similar sets of utterances. These experiments both illustrate the interaction of cognitive and cultural processes, as well as the ability to tease them apart.

The experiments conform to observations of systems of speech sounds in existing languages and the way they change: people have a tendency to apply the same processes (duplication and borrowing in the experiments; phonological
processes in real languages) in different utterances and to re-use existing building blocks. In addition, humans appear to apply similar processes to the building blocks themselves. In the experiments this is observed in the case of mirroring of building blocks, and in real languages it happens when one feature spreads over different phonemes or when utterances are re-analysed in terms of new sets of building blocks. In this way, whole sets of new speech sounds can appear (or disappear) in a language in a single operation. We can therefore tentatively conclude that complexity of phonological systems is due to cognitive mechanisms that re-use and generalize building blocks.

Due to the cognitive biases to learn and generalize complex sets of speech sounds, change of (systems of) speech sounds is not a completely random process. This does not alter the fact that language change can be viewed as a process of cultural evolution, just that the operations that introduce variation are not as random as those of biological evolution. This is a well-known potential difference between cultural evolution and biological evolution (e.g. Dennett, 2006) but it is one that is sometimes overlooked by linguists applying evolutionary theory to language change (e.g. Blevins, 2004). Experimental investigation of language creation, change and complexification as illustrated in this chapter, may help to clarify how variation is introduced and how complexity is created in language.

A discrepancy between the experiments and real language is that in the experiments rather radical changes happen quickly and frequently, whereas real languages are more stable over generations. This is most likely due to a much stronger conformity bias in the case of real language: language users are exposed to a much larger set of linguistic stimuli, and therefore learn much more accurately than in the
experiments. Moreover, language users cannot change their language too radically as otherwise they would be unable to communicate with others in the community.

The work described here is just the beginning of the experimental investigation of how culture and cognition interact in the emergence of phonological complexity, and a rather rough beginning at that. The work does form part of a movement to apply similar techniques to different aspects of language, most notably syntax and morphology, but also pragmatics and semantics (Kirby et al., 2008; Galantucci, 2009; Scott-Phillips & Kirby, 2010). There are a large number of issues that still need to be addressed, including: what is the effect of adding meaning to the tasks? What is the effect of interaction between speakers? Which phenomena are due to learning and reproduction and which phenomena are due to invention of utterances? Which processes are due to general cognition and which processes are due to language-specific cognition? Finally, there is the issue of the extent to which it is possible to exclude all influence from the linguistic knowledge that the participants of the experiments already have. All these questions notwithstanding, the approach of experimental cultural learning promises to be a useful new way to investigate old questions.

**Acknowledgements**

This work was financed by the NWO vidi grant 276-75-007 “modelling the evolution of speech”. The author wishes to thank Tessa Verhoef, Salikoko Mufwene and an anonymous reviewer for comments on the manuscript.
References


Griffiths, Thomas L., Kalish, Michael L., & Lewandowsky, Stephan (2008). Theoretical and empirical evidence for the impact of inductive biases on cultural
evolution. Philosophical Transactions of the Royal Society of London B, 363(1509), 3503-3514.


