INTERNATIONAL SCHOOL FOR ADVANCED STUDIES
Cognitive Neuroscience Curriculum

Academic Year 2014/2015

Constraints on Language Learning

Behavioral and neurocognitive studies with adults and children

Thesis submitted for the degree of
Doctor of Philosophy

Advisor:
dr Giosuè Baggio

Candidate:
Iga Nowak

4 February 2016

Contents

Chapter 1

Introduction

1.1 General Introduction

Language is fascinating and lies at the heart of what it is to be human. Our ability to express an unlimited range of thoughts and feelings through sounds, written signs, gestures, and the various other means of communication humans have invented, is unparalleled.

There is a long standing disagreement about the origins of language, both in phylogeny and in ontogeny, and what are the forces that shape it over time. A widely held belief is that all humans are endowed with a language-dedicated genetic program, which results in the 'growth' of a 'language organ' instantiated in the human brain (Chomsky, 1965). Because of this, so the theory goes, all human languages share a set of underlying principles or 'universals' (Greenberg, 1963, Chomsky, 1965). But this is not the only available explanation for the ubiquity and similarity of the structures inherent in human languages. More recent and alternative approach suggests that languages are shaped by the perceptual and cognitive constraints of learners (Kirby, 2001; Hurford, 2007; Christiansen & Chater, 2008; Culbertson, 2012). In other words, human biology and the human environment at large provide us with domain-general mechanisms that allow us to acquire various skills and abilities: language is just one of them.

The theory that language is genetically encoded relies on two main observations: the fact that all humans speak languages and that all languages appear to share highly specific features. One alternative theory emphasizes instead that language has originated and is being transmitted in the context of human social and cultural interactions. Empirical support for the latter view comes from our extraordinary capacity to learn from others, and from experiments suggesting that key features of language are shaped by the historical, cultural and pragmatic forces (Evans & Levinson, 2009; Dunn et al., 2011; Pagel et al., 2013). This view of language is consistent with the idea that the function of language shapes its structure.

Growing evidence suggests that the differences and similarities of the world's languages are neither the direct consequence of genetics nor of environment. Rather,they are constrained by both. If languages varied only very slightly from one another, then the idea that languages emerge principally from a common biological source would be fairly convincing. But if languages are shaped by communicative demands, diversity rather than similarity should be the hallmark of human languages. When we look at the world's languages, the amount of variation over space and time is startling. Evans & Levinson (2009) have shown in their seminal paper that humans are the only species with a communication system whose main characteristic is variation and not homogeneity. The differences between languages are greater than nativist accounts would predict. Yet all grammars are predicted to be similar in a few key and abstract characteristics (Greenberg, 1963). The idea that language is exclusively a product of nature or nurture is too simplistic.

The challenge for research at this stage is to identify the biological, cognitive and social constraints that have shaped languages, and determine how their interplay led to the structural and semantic patterns observed in natural languages. Pinpointing the relative effect of these constraints on language is one of the greatest modern challenges in the field. Christiansen & Chater (2008) identified four classes of constraints that interact in shaping languages: (1) perceptuo-motor factors, (2) cognitive limitations on learning, (3) constraints from thought and (4) pragmatic constraints. The latter class, pragmatic constraints, is known to be involved in linguistic communication, but for many years was overlooked by researchers.

A new trend has emerged in recent years in which language is treated as a dynamic population-based system, where speakers and listeners choose variants from a pool of linguistically plausible options in a way that is regulated by both social and cognitive constraints (Beckner et al., 2009; Hruschka et al., 2009). The implications are that language is an evolving system that is being shaped by its speakers.

In this novel framework, to understand language ontogeny and typological distribution, it is crucial to understand how languages are being acquired and how they change as they are being acquired. In general, if some aspects of linguistic structure arise from social or non-linguistic constraints, then it must be that linguistic structure in not exclusively determined by a language-specific ‘organ’ of the mind (Chater & Christiansen, 2010). Recent work (Hudson Kam & Newport, 2005; 2009; Culbertson et al., 2012) suggests that changes in linguistic systems arise from the individual processes of learning. However, as language is largely used in social interactions it is important to keep in mind that its structure is at some level dependent on its role in our social communication (Tomasello, 2008; Croft, 2009; Smith & Wonnacott; 2010).

This thesis will contribute to a body of experimental work addressing the question of whether language learning plays a role in certain fundamental design properties of natural languages. Methodologically, this thesis seeks to extend the artificial language learning paradigm, investigating whether learners are sensitive to the constraints embodied by key properties of languages. For example, we will explore whether communicative pressure influences the final outcome of language learning, namely how the structures that are acquired by individuals are transmitted to downstream generations. We will also explore how basic language learning constraints operate in different age groups and, importantly, cross-linguistically. Next to the behavioral experiments focusing on learning and its outcomes, we will look at preliminary electrophysiological correlates of basic compositional processing in the early stages of learning a miniature artificial language using electroencephalography (EEG). In this general introduction I will briefly discuss some of the relevant concepts and methods which will be used in three studies that constitute this thesis.

1.1.1 Artificial Language Learning

Data from language corpora, language acquisition and language change studies have provided valuable insights into language learning. One question, then, is why use artificial language learning paradigms instead of studying L1 and L2 learning? Firstly, given the complexities of natural languages and their acquisition, isolating particular features of language in order to study them is difficult. Artificial languages allow one to design a specific language with certain properties of interest while removing all other properties and controlling the linguistic data that participants are exposed to. This further allows one to study basic properties of languages such as compositionality, which is known to be much more difficult with natural languages due to the many empirical challenges to compositional semantics. Moreover, it is impossible to control for frequencies of words or of different structures in the input of L1 and L2. This creates potential confounds that should be taken into account (Izumi, 2003; Reali & Christiansen, 2007; Pagel et al., 2007). Artificial language learning research aims to simulate natural language learning and at the same time tries to minimize these confounds. Culbertson (2012) in her review provides extensive arguments as to why artificial languages allow us to explore hypotheses that would be otherwise impossible to test in the absence of controlled conditions.

Secondly, when focusing on language learning, it takes months or years for an L2 learner to reach a level of proficiency sufficient to enable communication, and even then the learning outcome is difficult to predict and impossible to properly control. This is where miniature artificial languages prove useful. They allow us to study language learning from novice to native-like proficiency within a short period of time. Miniature artificial languages resemble natural languages in that they are composed of a set of verbal items and a set of rules governing them. Like in natural languages, the set of rules can specify category membership, word order and other constraints on linguistic structure. This set of rules is limited, allowing us to study certain features of language and the learning situation in isolation from the complexity of natural languages.

Some researchers object to the relevance of conclusions drawn from artificial language studies as not being directly comparable and relevant to natural language learning. In this light, it is important to mention neuroimaging studies (more in Chapter 4) using artificial language that have shown that violations similar to those found in natural languages yield comparable ERP effects (Friederici et al., 2002) and recruit similar brain networks (Opitz & Friederici, 2003; Opitz & Friederici, 2004). One of the stronger criticisms of artificial language learning is the fact that adults possess a set of strategies they have acquired through years of experience as language users and the potential bias from their native language. However, so far there is no evidence confirming that adults in artificial language experiments adopt learning strategies transferred from linguistic conventions of their native language. The second criticism comes from the fact that a large proportion of artificial language studies have focused on adult learners. Some have argued that children bring different domain-general cognitive abilities to the language learning process and also learning mechanisms which adults do not employ.

We have directly addressed both criticisms in our behavioral studies. Firstly, addressing the second criticism, we have included two different age groups, preschool children and adults. To address the first criticism we have conducted our experiments with speakers from different language groups (Romance, Germanic & Slavic). Slobin and his colleagues (Berman & Slobin, 1994; Slobin, 1997) have claimed that cross-linguistic studies are more well suited to understanding putative universal features of language learning. The majority of the findings in cognitive linguistics are based on English speakers. English has unusual properties cross-linguistically that may render the results not entirely representative (e.g., the contrast between morphologically complex and simple forms). In two of our studies we focused on speakers of different languages: Italian, Polish and German. Patterns of language learning may vary across different linguistic communities. We believe that comparing different languages (in this case especially morphologically), for example Italian and Polish, represents a strong test of the generality of our results across languages.

1.2 Methodological paradigms and techniques

In this section I will briefly describe the experimental designs that allowed us to investigate artificial language learning.

1.2.1 Artificial Grammar Learning with Function and Content Words

In our first study we designed artificial grammars with a key property of natural languages: the distinction between more frequent and shorter words ('function words'), and less frequent and longer words ('content words'). This distinction is a universal feature of human languages: function words define and signal sentence structure, while content words convey meaning. This divide is present in all languages of the world. Function words contribute to language learning in two ways: they help to categorize content words (in English for example nouns are usually preceded by determiners) and indicate rules, thus increasing the learnability of structural rules. As a consequence, functions words act as anchors cueing to which structural role and position other constituents should be encoded. Overall, function words contribute to the higher learnability of linguistic structures (Morgan et al., 1987; Valian & Coulson, 1988). Therefore to bring artificial languages closer to natural languages and increase its learnability, for our study we constructed sets of pseudo-word strings with shorter and more frequent monosyllables, and longer and less frequent polysyllables, to mimic function and content words in natural languages.

In our experiments we used a string recognition task to evaluate the learners' performance. The task was composed of two phases. In the training phase participants listened to the strings generated by artificial languages and were simply instructed to pay attention. Afterwards, in the test phase, they were asked which of the strings (or string sets) they had heard belonged the to the language that they had been exposed to in the training phase. Participants then received feedback after each judgement, which, while not uncommon (Bahlmann et al., 2008) it is rather unusual in artificial language learning studies, as the role of feedback is controversial in language acquisition. Dale and Christiansen (2004) explored the role of feedback in artificial language learning studies and concluded that it may facilitate participants to reach a higher level of proficiency. Therefore, even if the learners in natural environments can acquire natural language without much feedback, we tested preschool children: in this case feedback, as a kind of reward, was necessary to keep them attentive and interested in the upcoming trials.

1.2.2 Iterated Signalling Games

In iterated learning, participants acquire a miniature artificial language and are subsequently tested on their knowledge. One of the definitions of iterated learning states that it is a 'process in which the behaviour of one individual is the product of observation of similar behaviour in another individual who acquired the behaviour the same way' (Scott-Phillips & Kirby, 2010 p. 411). However in the standard iterated learning paradigm this observation means that the learning output from the test phase of one individual is used as an input for the next participant. There is no coordination between participants. The flow of information is unidirectional without a social component inherent in language transmission.

In our experiments, described in Chapter 3, we introduced an interactional model of social learning - *signaling games*. Signaling games (Lewis 1969; Skyrms 2010) are a class of dynamic games of incomplete information in which a sender has private access to a selected state but the receiver does not. The sender sends out a signal, the receiver then acts upon it and chooses an appropriate action. If the chosen action matches the state, the game is a success for both players. A signal then is a product of the dynamic interaction between sender and receiver: the association between a signal and an action is decided jointly by the sender's and receiver's actions. Hence information flow is bidirectional.

We applied iterated signaling games in diffusion chains in an experimental setting. The linear diffusion chains we used involved fixed roles, and each generation consisted of only two individuals. By employing iterated signaling games we were able to study both vertical and horizontal transmission at the same time. One of our main aims was to provide a proof of concept for iterated signaling games, that is, to show they are a viable model of the cultural transmission of languages, complementary to Kirby et al.'s (2008) iterated learning.

1.2.3 ERPs and Language Studies

Event-related brain potentials (ERPs) reflect the on-line electrophysiological brain dynamics of cognitive processes with a high temporal resolution in the order of milliseconds. In the study reported in Chapter 4 we recorded electroencephalograms (EEGs) from participants one day after they learnt a miniature artificial language. The ERP technique has proven to be an important tool in exploring the neural architecture of language. However, very few studies have been conducted combining artificial language learning/processing and EEG. Combining EEG with behavioral measures increases the chance of identifying effects of language learning, as ERPs have been shown to be sensitive to the effects in language learning studies that were not found with behavioral measures (McLaughlin et al., 2004; Morgan-Short et al., 2012). The basic question in this type of research is whether the novel language (non-native) is processed in a similar way to native one. Many studies have reported differences in syntactic and phonological processing of non-native language (for a review see Mueller et al., 2005). However, some studies have reported indistinguishable processing of L2 reaching native-like patterns (Birdsong, 1992) highlighting the correlation between proficiency and certain syntactic ERP effects (Hahne et al., 2003; Friederici et al., 2002). The results in the lexical domain indicate more similarities than differences between native and non-native speakers. Overall, studies using miniature artificial languages are a promising start that could allow us to solve some remaining questions about the conditions necessary for learners to reach native-like processing.

1.3 Outline of the thesis

Following this Introduction, I will report the results of an experiment on artificial grammar learning in Chapter 2. We aimed at investigating whether learners more easily acquire patterns found in all natural languages, as opposed to unattested structural patterns. Here we used extended artificial language learning (more specifically, artificial grammar learning) to study language learning in two age groups (preschool children and adults) from three language groups (Italian, Polish and German).

Chapter 3 addresses the study on language change as the result of social coordination and language learning. In this experiment we used speakers of two languages (Italian and Polish) to study how properties of natural languages are shaped through vertical and horizontal transmission and whether our experimental model of iterated signaling games is a viable model to investigate language change.

Chapter 4 is an attempt to validate signaling games as away to study early stages of language learning, in particular using ERP measures to assess language processing of basic syntactic structures. We introduced a two-day experimental protocol: on the first day participants learnt a new language while on the second day they were tested on it, allowing for post-learning consolidation. We measured activity during the visual presentation of linguistic stimuli presented in either the correct or incorrect context, which varied in two main ways: violations of word order or violation of semantic content.

Finally in Chapter 5 I will summarize the main findings, discuss the broad conclusions which I would like to draw based on the work presented in previous chapters and point out some limitations and potential future directions.

Chapter 2

Artificial grammar learning in preschool children reflects universal linguistic patterns

Constituents occupy different positions within a sentence depending on how they are related, syntactically and semantically, to other constituents: this principle is known as structural order. No language seems to obey a linear order principle whereby constituents have fixed positions within a sentence. How does structural order constraints language learning? We exposed Italian-, Polish- and German-speaking preschool children to strings from artificial grammars where shorter and more frequent words ('function words'), and longer and less frequent words ('content words'), are placed according to structural, linear or free order. We found that children do not behave as unconstrained learners. In a test phase following acquisition, they showed a preference for strings whose structure they were exposed to during training, but significantly more so for structural and free order grammars than for linear order grammars. The effect was modulated by the children's first language. Our results indicate that structural and free order grammars may be easier to learn than linear order grammars. This was not found in Italian-, Polish- and German- speaking adults. These data are consistent with the existence of a developmental constraint on learning corresponding to a fundamental property shared by all natural languages.

2.1 Introduction

In recent years, research using natural and artificial languages has focused on learning constraints as a window into the power and scope of language acquisition. These learning constraints, according to classical arguments in linguistics and psychology, restrict the space of the hypotheses employed by language learners in inducing grammars, highlighting both capabilities and limitations of learners (Chomsky, 1965; Lightfoot, 1997). Grammars constrain the order in which constituents appear in phrases and sentences. These constraints apply to classes of constituents (e.g., articles and nouns), such that the occurrence of one constituent type depends on the presence and relative position of others (e.g., in English the placement of an article is determined by the position of a noun in a noun phrase). This principle is known as structural order. Constituents are organized in sentences following hierarchical organization which allows a combination of smaller units into more complex ones. Sentence structural dependencies are not dependant on the linear order of words. The order of words in sentences is thus not based on rules mapping the ordinal position of an element but by rules mapping the structural position of the elements in the sentence.

Natural languages are moreover organized into hierarchical frames. Words are bound together to form phrases, and phrases are in turn merged to form sentences. Language learners abstract grammatical structure from the linguistic stimuli they are exposed to in their environment. They can generalize beyond the given input, and the inference mechanisms they use to do so are likely to have inherent inductive biases for or against certain structural patterns. Learning biases, whether they are language-specific or domain-general, may arise out of early language acquisition (Saffran et al., 1996; Kuhl, 2004), they may reflect, albeit indirectly, the human biological make-up (Chomsky, 1965; Crain, 1991; Jackendoff, 2002), or they may be a combination of the two (Yang, 2004).

An example of learning constraints is indeed structural constituent order. On the surface, a sentence appears to be a string of words arranged in linear order, resulting from speech production, but is in fact organized hierarchically on different tiers. Syntactic and semantic operations are in general sensitive to a variety of structural relations between constituents. Although natural languages have largely structure-based rules, there exist linear position rules in specific domains such as phonology, and in certain cases also in syntax. In general, however, grammatical operations apply to words that play a particular functional role within a phrase or a sentence, rather than to items that occupy a certain position in a string (Pinker, 1984; Crain, 1991).

A classic example involves auxiliary-fronted interrogatives in English. Children never turn a statement like 'The man who is tall is Sam' into the ill-formed question 'Is the man who tall is Sam?' (Gomez & Gerken, 2000; Ambridge et al., 2008). Here, children are not applying a linear order rule, like 'move the first verb to the front of the sentence', but they use a structural constraint in the transformation, i.e., 'move the main verb to the front'. It suggests that children while acquiring L1 are already sensitive to structural relationships between grammatical entities like words, phrases and clauses, rather than to linear rules that may describe the surface sequence of words.

Structural order constraints on learning have been investigated using computational methods (Perfors et al., 2011), but experimental work at this level of generality is still lacking. Do structural order constraints exist in the mind of learners and, if so, can they be probed experimentally? We conducted a cross-linguistic series of behavioral experiments assessing the performance of preschool children and adults in learning and recognizing legal strings from simple artificial grammars featuring the distinction between shorter and more frequent words ('function words') and longer and less frequent words ('content words.') Our artificial grammars were constructed either based on plausible rules on the placement of function and content words in a string (i.e., structural order or free order), or based on implausible rules (linear order). All grammars used in our study were finite-state (Figure 2.1).

We tested children and adults from three different language groups (Romance, Germanic and Slavic) to assess the dependence of structural order constraints on the first language (L1) of learners. Languages make use of word order, inflection and prosody to link constituents in sentences (Morgan et al., 1987), but certain languages rely more heavily on one or another grammatical device for particular functional purposes. A notable difference is the amount of constituent order flexibility they allow. For example, in English changing constituent order also changes the relations between constituents. On the other hand, in inflectional languages, such as Polish, word order is freer, depending largely on the context. Dryer (1992) noted that some languages, like German, are treated as lacking a dominant constituent order, as different word orders are conditioned by syntactic factors. Also Italian allows all possible orders of subject, verb and object, depending on the context, with SVO still being predominant.

Both Italian and Polish are so called pro-drop languages, where subject is included only if it is new in the context. German requires a free-standing subject in declarative sentences. This results in verbs being often the first element in the sentences and verbs constituting a much higher proportion of content words in Italian and Polish compared with German. Italian inflectional morphology is reduced with contrast to Polish one. Polish content words are overtly morpho-phonologically complex, realised within a rich inflectional paradigm. Conjugation tables of Polish verbs provide over a hundred types of conjugational paradigms (Saloni, 2007). Nouns in Italian are not morphologically inflected for case (excluding pronouns), thus grammatical role of the noun is informative about the syntactic function of the word in the grammatical structure. Overall, we believe that these languages (Italian, Polish and German) display enough differences (and similarities), to allow us to investigate whether learning constraints are modulated by the first language of learners.

In recent years, the artificial language learning paradigm has become a driving force behind connecting typological universals with learning biases. Although studies of morphology or syntax are relatively rare, several studies have attempted to address learners' constraints in learning unattested syntactic patterns (Culbertson, 2012). Christiansen (2002) investigated the emergence of word order universals and suggested that placement of heads relative to their complement phrases can be explained as processing constraints. St. Clair et al. (2009) explored the learning bias in favour of suffixing over prefixing using artificial languages, providing further support that learning constraints reflect patterns found in cross-linguistic typology. However, the experimental work of learning constraint on elementary constituent order is still missing.

The ability to acquire linear rules has been demonstrated by several studies (Reber, 1989; Gomez & Gerken, 1999), while acquiring structural syntactic dependencies has provided conflicting evidence in artificial language studies. Perruchet and Rey (2005) and later de Vries and colleagues (2008) suggested that results showing the learning of syntactic structures could be explained by different mechanisms such as counting. Their results caused a debate about the learnability of structural dependencies as well as the need to study other then learning of embedded hierarchical structures. The interest of comparing learning of the two types of rules comes from the idea that languages which violate typological universals are predicted to be more difficult to learn and less likely to be acquired by new generations (Culbertson et al., 2012). Hence, learners should find it easier to extract the rules based on the structural representation of the presented stimuli. We focused on learning of two types of constituent orders: linear order that is based on the position of the elements in the linear sequence and structural order that is based on the position of the elements in the constituent structure. The linear order allows to investigate the learning of adjacent and non-adjacent dependencies based on the fixed position of the content and function words in the strings. The structural order allows to test learning of structures based on the relative position of function and content words in the strings of the artificial language. In our study, linear order determines the position of function words with respect to their ordinal position in the strings (e.g. the function words appear always and only in the first an last position of the string). While structural order focuses more on the relations between the two classes of words (the presence of function words depends on the content words). There is an ongoing debate as to what kind of rules should be used to study constituent order in artificial grammar learning experiments. Most studies focused on centre-embedding grammars (Hauser et al., 2002), however it is possible that learners apply simpler strategies (e.g. counting the elements) while learning these rules. Therefore, these types of grammars may not be the best choice to study learning constraints of constituent order. We aimed to introduce a type of artificial grammar learning rules which are closer to what is found in natural languages (mirroring the distinction between content and function words).

The complexity of natural languages makes it difficult to isolate the factors that play a primary role in language acquisition. In recent years, AGL paradigms have been used to avoid or check confounding factors on learning, such as prior knowledge of natural language grammars, as well as to control the input that learners are exposed to (Folia et al., 2010). A number of recent studies have shown that AGL learning by adults and children reveals biases that parallel with the typological universals found across languages (Singleton and Newport, 2004; Hudson Kam & Newport, 2009; Fedzechkina et al., 2012; Culbertson et al., 2012; for a review see Culbertson, 2012). Friederici et al. (2002) demonstrated that adults who learned miniature artificial languages display a similar real-time pattern of brain activation when processing language as native speakers.

However, some have argued that natural language input, enabling us to acquire syntax, includes not only bare strings of words – as is often the case in AGL experiments

– but also information on how words may group into phrases (Braine, 1966; Morgan & Newport, 1981; Valian & Coulson, 1988; Valian & Levitt, 1996; Gomez, 2002), based on cues such as prosody, the distribution of function words, and morphology. Morgan et al. (1987) tested whether such cues are necessary for successful acquisition of syntax in adults, and found that both prosodic cues and frequent function words improved the acquisition of syntax.

Here, we used artificial grammars with a key property of natural languages: the distinction between more frequent and shorter words ('function words'), and less frequent and longer words ('content words'). In natural languages, content words contribute meaning to a sentence, they are often stressed in speech, they tend to be longer and less frequent, and to have clear referents (Grimshaw, 1981). In contrast, function words have linking and structural roles, and determine a sentence's logical form. Function words are morphemes that are often associated with a particular constituent. For example, in English the prediction is that children will use very high-frequency morphemes like 'the' as anchor points, and observe what words co-occur with them (Valian & Coulson, 1988). Function words possess several characteristics that allow them to be easily identified by native learners: they are short (usually monosyllabic) and unstressed. In each language, the number of function words is typically small, and many occur with high frequency. Green (1979) found that language structure was unlearnable if no frequent function words were available.

Learning an artificial language with high frequency markers was easier than learning a structurally identical language where the frequency of the markers was lower (Morgan et al., 1987; Valian & Coulson, 1988). More recent work (Bell et al., 2009; Gervain et al., 2013) suggests that the presence of function and content words in training sets results in greater learnability of linguistic structures. Moreover, already 17-month old infants use word frequency as a cue for identifying function words (Hochmann et al., 2010), thus function words facilitate word learning by providing syntactic cues that used for learning content words. While different cues have been found to enable to distinguish and classify between content and function words (like prosody, semantics, age of acquisition), it is well established that function words are short and their frequency decreases with their length, while content words are longer and their frequency is independent of their length (Miller et al., 1958). For our own AGL study, we constructed sets of pseudo-word strings with shorter and more frequent monosyllables, and longer and less frequent polysyllables, to mimic function and content words in natural languages.

Artificial Grammar Learning studies have explored two language learning mechanisms: statistical learning that plays a role in tracking linear order on the basis of transitional probabilities and mechanism essential for grammar-like learning of structures. Since it has been shown that already infants are able to extract statistical computations from artificial language input (Saffran et al., 1996), it has been further proposed that this mechanism may be powerful enough to account for language acquisition (Seidenberg, 1997). Yet, a separate subfield of artificial grammar studies have proposed that in order to fully acquire language, statistical learning mechanism is not enough and cannot fully explain the outcome of learning. Pena et al. (2002) suggested that the additional learning rule-like mechanism guides the nature of learning. Their main argument comes from the fact that when the test material differs from the training material, learners cannot just rely on transitional probabilities over the specific items present during the training. Thus, a mechanism that allows to abstract beyond specific items is needed. While statistical learning mechanisms have been primarily studied within the experimental framework of segmentation where participants have to extract smaller units from continuous stream, the mechanisms dedicated to structure learning have been studied within the framework of syntax learning studies where participants have to detect and generalize syntactic rules from strings generated from an artificial grammar. As the grammatical structure is usually held constant, while the vocabulary changes, it indicates that learners are able to abstract the grammatical structure beyond specific elements.

Overall, our AGL paradigm has the following characteristics (more details in Methods): spoken auditory stimuli (many AGL experiments use visual presentation of stimuli, which may be argued to tap into different memory processes than speech comprehension – see Culbertson et al., 2012); an implicit learning task in the training phase (Reber, 1989; Perruchet & Pacton, 2006; Rohrmeier et al., 2012); a 2-alternative forced-choice (2AFC) task in the test phase, which has been argued to be most suitable to assess children's language learning (Crain and Thornton, 2000); and generalization to strings with the same structure as strings from the learning phase, but with novel vocabulary.

In the training phase participants were presented with strings generated by one of the rules and in the test phase they performed a grammaticality judgment task where they had to indicate which strings were generated by the rule. We were thus able to study lexical generalization - when grammatical strings had different lexicon compared to those in the training phase but represented by the same rule. Learner's ability to discriminate grammatical from ungrammatical strings, despite the new words, was taken as evidence that they extract grammatical structures from the string sets (e.g. Gomez & Gerken, 1999; Finley & Badecker, 2009).

It has been suggested that there are cognitive and biological factors at play early in life ensuring language acquisition, and that these weaken as maturation progresses. In particular, one hypothesis is that constraints on learning that are specific to language acquisition undergo maturational decay (Newport, 1990). A second hypothesis is that language acquisition abilities decline due to the expansion of non-linguistic cognitive abilities (Ramscar & Gitcho, 2007), and as a result of effort in learning syntactic categories (Finn et al., 2014). It remains to be seen how strong the learning constraints are for different age groups, and whether in some cases they will function as absolute constraints. It is possible that adults learning artificial languages are able to overcome, to some extent, constraints that are effectively absolute in language learning.

To assess whether structural dependencies constrain learning in different ways at different stages of cognitive and biological maturation, we conducted the same experiments with children and adults, in both cases with Italian, Polish and German native speakers.

2.1.1 Aims

We aim to address a theoretically important issue on the nature of learning linguistic structure, namely whether learners more easily acquire patterns found in all natural languages, as opposed to less plausible structural patterns. We used an AGL paradigm to study whether preschool children and adults from three different language families have an implicit inductive bias, such that they find it easier to learn a structural order or a free order grammar (as evidenced by recognition performance in a test phase) as compared to a linear order grammar. If there were no constraints on learning an artificial language, everything else being equal, participants should learn the two grammars equally easily. That is our null hypothesis.

2.2 Experiments 1-2

2.2.1 Methods

2.2.1.1 Participants

In Experiment 1, the sample consisted of 44 children (mean age: 4.5; age range: 3.2-5.9, 23 female): 22 children were Italian and 22 were Austrian. In Experiment 2, the sample consisted of 56 children (mean age: 4.8; age range: 3 - 5.7, 32 female): 28 children were Italian and 28 Polish. The subjects across groups were matched with respect to their age.The age was chosen due to linguistic development and maturational factors. Chapman & Kohn (1978) reported that only at this age children are able to process sequences with three phrases. Moreover, Slobin & Bever (1982) found that children within the age 2 to 4 from four different language groups were prepared to learn both inflectional and word-order artificial languages. In the first experiment, 5 children, and 8 in the second, were discarded for failure to attend to or understand the task. Children were recruited from Italian, Austrian and Polish kindergartens, and participated on a voluntary basis. They were given a candy at the end of the session. Parents of children had been informed about the experiment. Only the children of parents who returned a signed consent form took part in the study. The study was approved by the Ethics Committee at the International School for Advanced Studies in Trieste.

2.2.1.2 Stimuli

We constructed pseudo-word strings containing shorter, more frequent monosyllables, mimicking function words, and longer (2-3 syllables), less frequent polysyllables, mimicking content words. Pseudo-words were consistent with Italian, German and Polish phonotactics, but they were not meaningful words in either of the languages. In Experiment 1, children from Group 1 (structural order, SO) listened to strings obeying a structural order rule that each function word follows a content word, as in consistent left-branching or head-final languages (e.g., Japanese), where the head of the phrase follows its complements. That is, in a phrase, a function word is preceded by a content word, thus placing function words hierarchically in a different position with respect to content words. Children from Group 2 (linear order 1, LO1) were exposed to strings

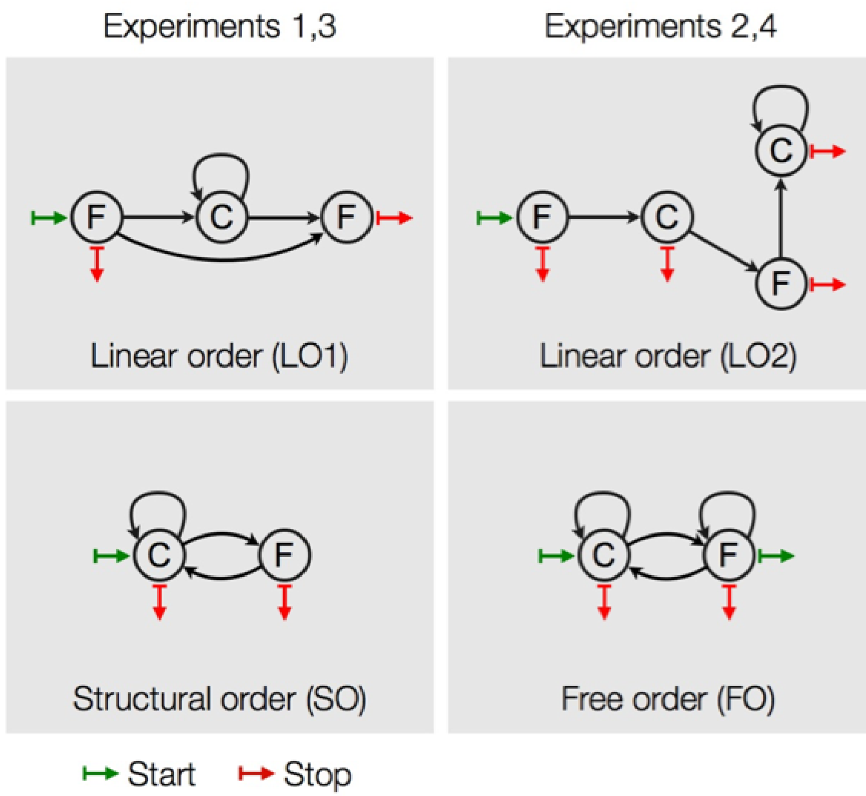


Figure 2.1: State diagrams constraining the generation of grammatical strings in the finite state grammars used in experiments 1-4. Black arrows represent legal transitions between symbols, i.e., function (F) and content (C) words; green and red arrows show the possible initial and final states of each generation process, respectively.

following a typologically less plausible linear order principle that all strings must begin and end with a function word (Figures 2.1-2). Here, there was no hierarchical or structural relation between function and content words. Function words were placed in the sentences based only on the linear rule. Even though it is easy to devise examples from natural languages where sentences begin and finish with function words, their occurrence is never constrained by a rule on the position they should occupy.

One may argue that SO uses only local dependencies, and as such it may be easier to learn than an LO grammar built on an implausible linear rule, that could also be seen as relying on a long-distance relation between the first and last function words. In free order languages, one can acquire grammatical structures only by representing non-adjacent dependencies. Long-distance dependencies are critical in that learners

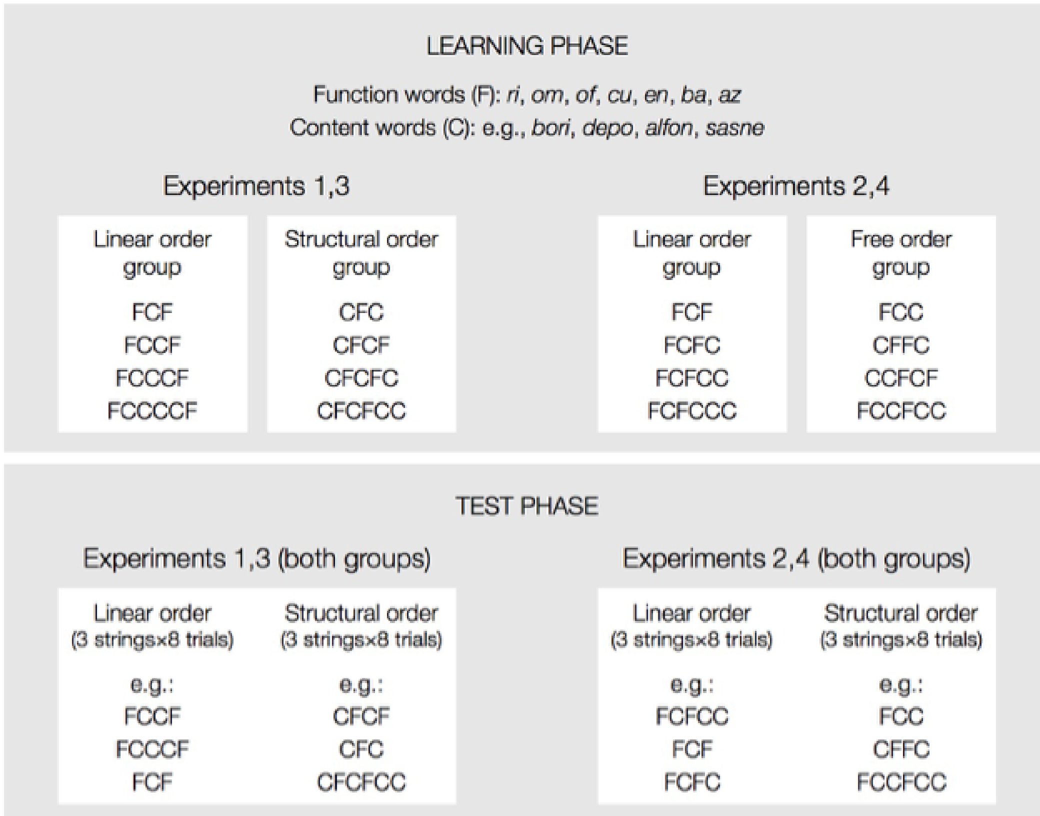


Figure 2.2: Summary of the experimental design and testing protocol. Abbreviations: C, content word; F, function word.

must coordinate elements that are far apart in a sequence. Thus, in Experiment 2, instead of structural order, we employed free order (FO) patterns where strings lack a governing rule on the position of content and function words, indeed as in free word order languages. In the FO grammar, there is no governing rule, but it may be still easier to recognize FO strings, as compared to LO strings based on a linguistically implausible principle. Children in Group 1 (FO) listened to free order strings, and children in Group 2 (LO2) were exposed to strings following an implausible linear order rule that function words appear in the first and third position in the string (Figures 2.1 and 2.2).

Each of the language had four possible sentence types: ranging in length from three to six words. The vocabulary of the language consisted of monosyllabic pseudo-words (for example: ri, om, en) and longer pseudo-words (for example: bori, sasne).

In total, each function word appeared six times more frequently (Valian & Coulson, 1988) than each content word. In the training phase, each content word was presented once; each function word appeared every time in a different position and in a different sentence type. For the test phase, we used new pseudo-words that learners were not exposed to before. Crucially, children from the two grammar groups (SO and LO1, or FO and LO2) were exposed to strings containing the exact same pseudo-words, varying only in the order in which they were arranged, depending on either the structural order (SO), the free order (FO), or the linear order (LO1, LO2) grammars they were assigned and exposed to during the learning period.

Our grammars used exactly the same pseudo-words, and the only difference was in the structure: how different word types were distributed within a string. Strings were matched across grammars for length, number of pseudo-words, number of function and content words (with the exception of 3-word strings). The complexity of pairs of n-pseudo-word structures across grammars (e.g., for n=5, FCCCF and CFCFC in Experiment 1; Figure 2) was matched using Kolmogorov complexity (Lempel & Ziv, 1976; Kaspar & Schuster, 1987) and Shannon entropy (Shannon, 1948). These measures can vary considerably for shorter binary strings. We computed the complexity of binary strings of Fs and Cs of the same length, and not of actual strings of the same length, in which Fs and Cs are replaced with pseudo-words as these will trivially have the same complexity. In our stimulus set no function or content word is repeated within a string.

For the same reason, n-gram frequency will also be trivially matched. In a further analysis, we calculated the Kolmogorov complexity (K) and Shannon entropy (E) for structures (i.e., again, binary strings of Fs and Cs) up to strings of length 12 (exceeding the maximum length of actual experimental strings, i.e., 6), to determine whether the complexity of strings from two grammars would diverge with increasing string length. As a further guarantee on the absence of structural design confounds in our stimuli, we found that was not the case: comparing the mean of K and E across strings between grammars (i.e., linear vs structural in Experiment 1; linear vs free in Experiment 2) using Wilcoxon rank sum tests with continuity correction, no significant effects were observed (Experiment 1, K: $W=29.5$, $p=0.129$, E: $W=42$, $p=0.568$; Experiment 2, K: $W=39.5$, $p=0.447$, E: $W=49$, $p=0.97$). These analyses show that string pairs of a certain length have exactly the same structural complexity across the grammars being compared in each experiment. This desirable property scales up to strings of increasing length, including strings that were not used in the present experiment.

We recorded a trained female phonetician reading aloud the strings in a natural animated voice. Recordings were preferred to text-to-speech software to ensure that stimuli were engaging enough to attract children's attention throughout the experiment. As grammars differed only in how pseudo-words were arranged, the same acoustic token for each word was used in strings from both grammars. This ensured that children could not use idiosyncrasies in the pronunciation of the pseudo-words in either grammar to discriminate grammatical from ungrammatical strings. We have not used phrasal prosody, as suggested by Valian & Levitt (1996), who investigated how different structural cues – frequency, reference field and prosody – interact in syntax learning. Their results suggest that learners attend to prosody as a cue only when no other cues are available. In our experiment, we used highly frequent (function) words which acted as markers, thus prosody would enhance learning only in a limited way. Strings were digitized and edited with Audacity 1.3.14. Audio files were normalized to a mean intensity of 60 dB. The strings were separated by 1s of silence when they were presented to children during the learning phase, and pseudo-words within a string were separated by 0.2 s of silence (Marchetto & Bonatti, 2013).

2.2.1.3 Apparatus

To heighten children's interest in the stimuli and task, we used a colorful puppet theatre and two cloth puppets (Crain & Nakayama, 1987). The stimuli were presented using loudspeakers invisible to children. The puppets were chosen to be attractive enough to keep children attentive. The theatre (about 1 m wide) was set on a table. The experimenter stood behind the theatre, invisible to children during presentation of the stimuli to prevent them from getting distracted. In the test phase, two boxes with lids were placed in front of the loudspeakers. Children were tested individually in a quiet room and in normal lighting conditions at their kindergarten.

2.2.1.4 Procedure

The experiment consisted of learning phase, followed by a test phase. We used a between-subjects design. Children were randomly assigned to one of two counter-balanced grammar conditions: structural order (SO) or linear order (LO1) in Experiment 1, and free order (FO) or linear order (LO2) in Experiment 2. Children were exposed for approximately 5 min to a subset of all grammatical strings from the grammar they were assigned to, and they were subsequently tested for their ability to generalize to strings that had the same grammatical structure as in the training phase but using novel vocabulary items (i.e., what is commonly referred to as a 'transfer test'). The ability to recognize grammatical structure despite a new vocabulary may be taken as evidence that learners have abstracted essential aspects of the grammatical structure of the strings from the input (Gomez et al., 2000).

Learning phase - Children were introduced to puppet A and were told he was from a distant land and spoke a different language. They were asked to listen carefully to puppet A as they would be asked questions about his language later on. Children were then exposed to 48 naturally recorded strings. Each sentence was repeated three times, resulting in a total number of 144. The learning phase lasted approximately 5 minutes during which children were looking at puppet A on the stage while listening to the auditory stimuli coming from both loudspeakers simultaneously. The learning phase was immediately followed by the test phase.

Test phase - In the test phase we used a 2AFC task to assess children's ability to recognize strings with a similar structure to strings they heard during the acquisition phase. Two boxes were put on the left and right sides of the stage, so that children could see them. Behind each box, on the back of the stage, lay the two loudspeakers. Children were introduced to puppet B. They were told that puppet B speaks a different language than puppet A, and that now both puppets will play a game with them: they will hide inside the boxes, one in each box, and they (the children) will have to listen to sentences coming from each of the boxes. Children listened to a sequence of three strings (i.e., a trial) coming from one box (first loudspeaker), all of which were constructed based on the order rule of the grammar they were exposed to during the acquisition phase, and to three strings coming from the other box (second loudspeaker), all of which were constructed according to the rule of the grammar they were not exposed to (e.g., LO1, if they were exposed to SO in the test phase, or vice versa; LO2, if they were exposed to FO, or vice versa). Associations between string from a grammar (exposed/non-exposed to during the learning phase) and boxes or loudspeakers (left/right) were randomized.

The experimenter asked the child in which box puppet A was hiding, and the child responded by pointing to or verbally referring to one of the boxes: no other type of answer was considered valid. The experimenter opened the chosen box and gave the child feedback, showing whether puppet A was in the box or not. Next, the other box was also opened, and its content was shown to the child. This test procedure was repeated on 8 consecutive trials (with each trial consisting of the presentation of 3 strings) and lasted about 12”. We chose to administer 8 trials for two reasons. Firstly, increasing the amount of trials would potentially lead to children’s loss of concentration and interest, thus altering the results. Secondly, the most informative data are likely to be obtained immediately after training, in the first few trials of the test phase: because our test trials include feedback, employing more trials would eventually bring performance close to ceiling level, decreasing the chances of seeing differences between conditions as the test phase becomes more extended.

2.2.1.5 Data analysis

We counted the number of correct trials for each child in each group, and we compared the means relative to chance level (4) using one-sample Wilcoxon signed-rank test (the data in each group was tested for normal distribution (Kolmogorov-Smirnov Test), in some groups the data was not normally distributed hence we opted for the non-parametric tests). We moreover defined a ‘learning index’ as the difference in the number of correct responses between the second (trials 5-8) and the first half (1-4) of the test phase, which were compared by means of Wilcoxon signed-rank test. This simple measure of the increment in correct responses brought about by learning is essentially equivalent to computing the slope of the rolling mean of correct responses over all 8 trials. Finally, we entered all data in two ANOVA models using the between-subjects factors Grammar (4 levels: LO1, LO2, SO, FO) and L1 (3 levels: Italian, German, Polish). The dependent variables were either correct responses or values of the learning index.

2.2.2 Results and Discussion

The results of Experiments 1 and 2 are shown in Figure 2.3 and Figure 2.4. In Experiment 1, children in the SO group performed better (M=5.14 correct trials, with 8 as the maximum) than children in the LO1 group (M=3.77). The performance of children in the SO group only was above chance (Table 2.1). A similar pattern was found in

Experiment 2 (Figure 2.3, Table 2.1). Children performed better in the FO condition (M=4.89) than in LO2 (M=3.89), and their performance was significantly above chance level only in FO (Table 2.1). The ANOVA reveals a significant main effect of Grammar ($F(3,92)=4.498$, $p=0.005$). There was no main effect of L1 ($F(2,92)=1.29$, $p=0.28$) and no interaction between the factors Grammar and L1 ($F(2,92)=2.278$, $p=0.108$).

Figure 2.3(b) shows the effect of learning during the test phase. Here, the learning index is the difference in the number of correct responses between the second (trials 5-8) and the first half (1-4) of the test phase. In Experiment 1, children showed a larger learning index in SO than in LO1, and similarly, in Experiment 2, the learning index is larger in FO than in LO2. The ANOVA shows an effect of Grammar ($F(3,92)=4.832$, $p=0.004$) and of L1 ($F(2,92)=4.772$, $p=0.011$) but no interaction between Grammar and L1 ($F(2,92)=0.610$, $p=0.545$).

Figure 2.4 shows the results of Experiments 1 and 2 for each language group separately. A similar trend can be observed across groups, favoring structural and free order over linear order. However, Italian children did not perform like German and Polish children: these were the only two groups that performed above chance (Table 2.1). In Experiment 1, Italian children were slightly more proficient at learning, during test, strings from the LO than from the SO grammar, unlike children from the other language groups (Figure 2.4b). Presumably the feedback provided during test had a more marked effect on recognizing LO than SO strings.

Performance for LO in the German group is significantly below chance: German children chose more often SO strings though they were exposed to a LO grammar. We found a similar, yet stronger, trend for Polish children. During the test phase, they were seemingly ‘unlearning’ LO (i.e., the learning index here is negative) or they were becoming more attracted to FO. Performance was significantly above chance only in the FO group.

These differences between L1 groups in test-phase learning effects are consistent with the observed main effect of L1 in the ANOVA (see above). These data show there are constraints operating during learning, either in the form of a preference for SO/FO, or a dis-preference for LO. Furthermore, the effect of these constraints on learning is modulated by children’s L1.

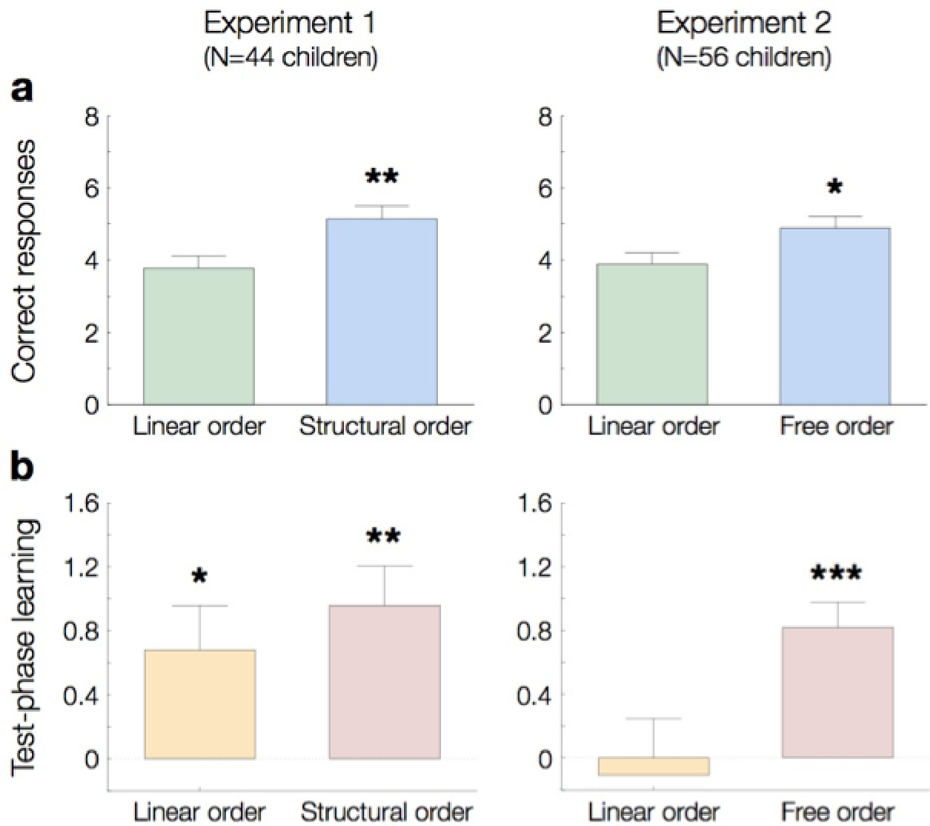


Figure 2.3: Results of Experiment 1 and Experiment 2. Panel (a) shows the average number of correct responses over all 8 trials. Panel (b) shows the learning effect calculated as the difference between correct responses in the second and the first half of the test phase. Error bars denote standard errors. Asterisks indicate statistically significant effects (Table 2.1) in one-sample Wilcoxon signed-rank test relative to chance (correct responses, panel a) and Wilcoxon signed-rank test comparing the second and the first half of the test phase (learning index, panel b). Signficance codes: $p < 0.05^{*}$; $p < 0.01^{**}$; $p < 0.001^{***}$.

2.3 Experiments 3-4: A comparison between adults and children

Potentially different language learning strategies in adults and children have attracted considerable attention (Hudson Kam & Newport, 2009; Fava et al. 2011; Finn et al., 2014). In order to study the developmental continuity of learning constraints, we tested

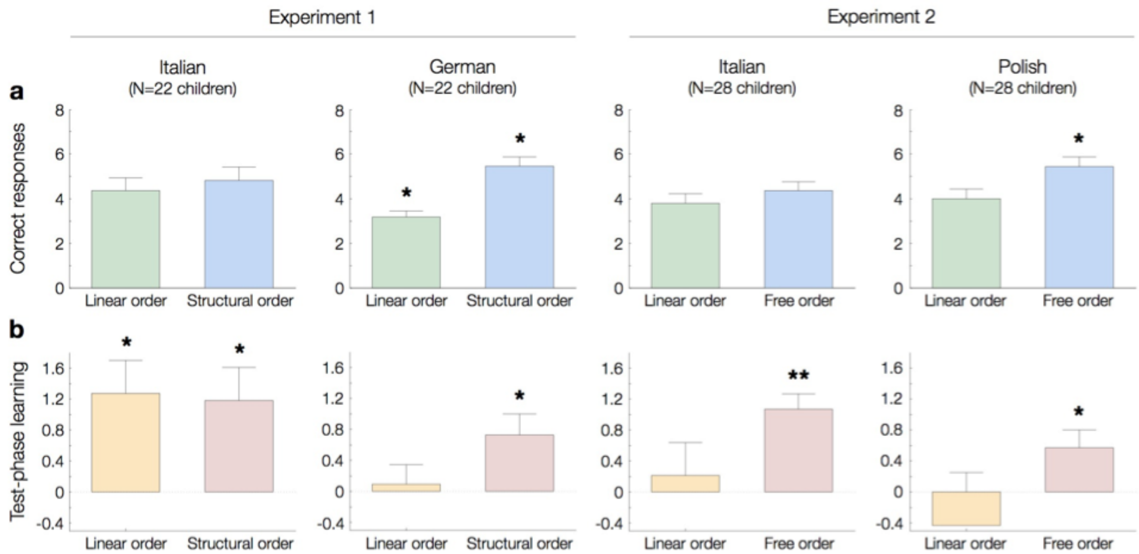


Figure 2.4: Results of Experiment 1 and Experiment 2 for Italian 1 (SO/LO1), German (SO/LO1), Italian 2 (FO/LO2) and Polish (FO/LO2) language groups. Error bars denote standard errors. Asterisks indicate statistically significant effects (Table 1) in one-sample Wilcoxon signed-rank test relative to chance (correct responses, panel a) and Wilcoxon signed-rank test comparing the second and the first half of the test phase (learning index, panel b). Significance codes: $p< 0.05^*$; $p< 0.01^{**}$; $p< 0.001^{***}$.

adults using the exact same experimental paradigm, and with the same stimuli and task used with kindergarten children in Experiments 1-2. To investigate how knowledge of a native language interacts with learning, when learners are confronted with an artificial language, we tested native speakers of three languages: Italian, German and Polish.

2.3.1 Methods

2.3.1.1 Participants

In Experiment 3, the sample consisted of 42 adults: 22 were Italian, 20 were Austrian (29 females, mean age 22.44 years, age range 19-35 years). In Experiment 4, the sample consisted of 42 adults (31 females, mean age 20.66 years, age range 19-27): 22 were Italian, 20 were Polish. Written consent was obtained from all participants, who were paid for taking part in the study.

		SO/FO Correct responses	LO1/LO2 Correct responses	SO/FO Test-phase learning	LO1/LO2 Test-phase learning
Experiment 1	M (SE) Z p r	5.14 (0.362) 2.63 0.009 0.40	3.77(0.335) 0.322 0.747 0.05	0.96 (0.25) 2.89 0.004 0.43	0.68 (0.274) 2.05 0.04 0.31
Experiment 2	M (SE) Z p r	4.89 (0.306) 2.54 0.011 0.34	3.89(0.306) 0.437 0.662 0.06	0.82 (0.155) 3.69 0 0.49	-0.11 (0.25) 0.662 0.508 0.01
Italian (Exp. 1)	M (SE) Z p r	4.82 (0.6) 1.32 0.185 0.28	4.36(0.576) 1 0.314 0.21	1.18 (0.423) 2.08 0.04 0.44	1.27 (0.428) 2.18 0.03 0.45
German (Exp. 1)	M (SE) Z p r	5.46 (0.413) 2.46 0.014 0.52	3.18(0.263) 2.25 0.024 0.48	0.73 (0.273) 2.13 0.03 0.45	0.09 (0.251) 0.378 0.705 0.08
Italian (Exp. 2)	M (SE) Z p r	4.36 (0.401) 0.82 0.413 0.15	3.78(0.447) 0.5 0.617 0.09	1.07 (0.195) 3.04 0.002 0.57	0.21 (0.422) 0.424 0.672 0.08
Polish (Exp. 2)	M (SE) Z p r	5.43 (0.429) 2.57 0.01 0.49	4 (0.432) 0.06 0.952 0.01	0.57 (0.228) 2.13 0.03 0.40	-0.43 (0.25) 1.61 0.107 0.3

Table 2.1: XYZ Mean (M), standard error (SE) and effect size (r). For structural order (SO), free order (FO) and linear order (LO) responses, the statistics are one-sample Wilcoxon signed rank tests relative to chance level (4). For the learning effect, statistics are Wilcoxon signed-rank tests comparing the first and second half of the test phase.

2.3.1.2 Stimuli, Apparatus, Procedure and Analysis

The same as in Experiments 1 and 2.

2.3.2 Results and discussion

Figure 2.5 and table 2.1 report the results of Experiments 3-4. In general, adults performed better than children. Their responses were significantly different only in the LO2 group (Mann-Whitney test comparing children's and adults' performance: SO: U=271, z=0.995, p=0.32; LO1: U=294.5, z=1.58, p=0.114; FO: U=249.5, z=0.918, p=0.359; LO2: U=459.5, z=3.4, p=0.001). There was moreover a significant difference in the learning index in the LO groups (Mann-Whitney test comparing children's and adults' learning during test: SO: U=243.5, z=0.314, p=0.753; LO1: U=312.5, z=2.05, p=0.04; FO: U=262, z=0.676, p=0.499; LO2: U=466.5, z=3.6 p=0.001; Figure 5b): adults from all L1 groups were getting better during test in recognizing LO strings, whereas children's performance during test improved for SO and FO grammars, showing an 'unlearning' or an FO-attraction effect for LO2 grammar in Experiment 2. In Experiment 3, both grammar groups performed above chance, and there was no significant difference in performance (U=241, z=0.53, p=0.596), suggesting that it was equally easy for participants in both groups to recognize SO and LO1 strings. However, in Experiment 4, only LO2 participants performed above chance, whereas in the FO group performance was at chance level. There was a significant difference in performance between LO2 and FO groups (U=317.5, z=2.495, p=0.013).

These results indicate that adults generalized from the training phase more easily when they were exposed to a LO grammar, featuring a consistent generation rule, compared to grammar without any apparent rule (FO). To test the hypothesis that Grammar and L1 had an effect on behavior, a between-groups ANOVA was performed, showing an effect of Grammar (F(3,76)=5.782, p=0.001) and L1 (F(2,76)=5.438, p=0.009), but no interaction between Grammar and L1 (F(2,76)=0.7, p=0.5). Figure 5b shows the effect of learning during the test phase. In both experiments, participants showed a larger learning index in the LO than in the SO/FO groups, especially so in Experiment 4. The ANOVA shows no effect of either L1 (F(2,77)=0.123, p=0.884) or Grammar (F(3,77)=1.703, p=0.173), nor an interaction of the two factors (F(2,77)=0.146, p=0.865).

Figure 2.6 and Table 2.1 show the results of Experiments 3-4 in each language group separately. All groups performed above chance, with the exception of the Italian FO group (Figure 2.6 (a); Table 2.2). In Experiment 3, performance was not significantly different between the two L1 groups. In all groups, participants were better at learning

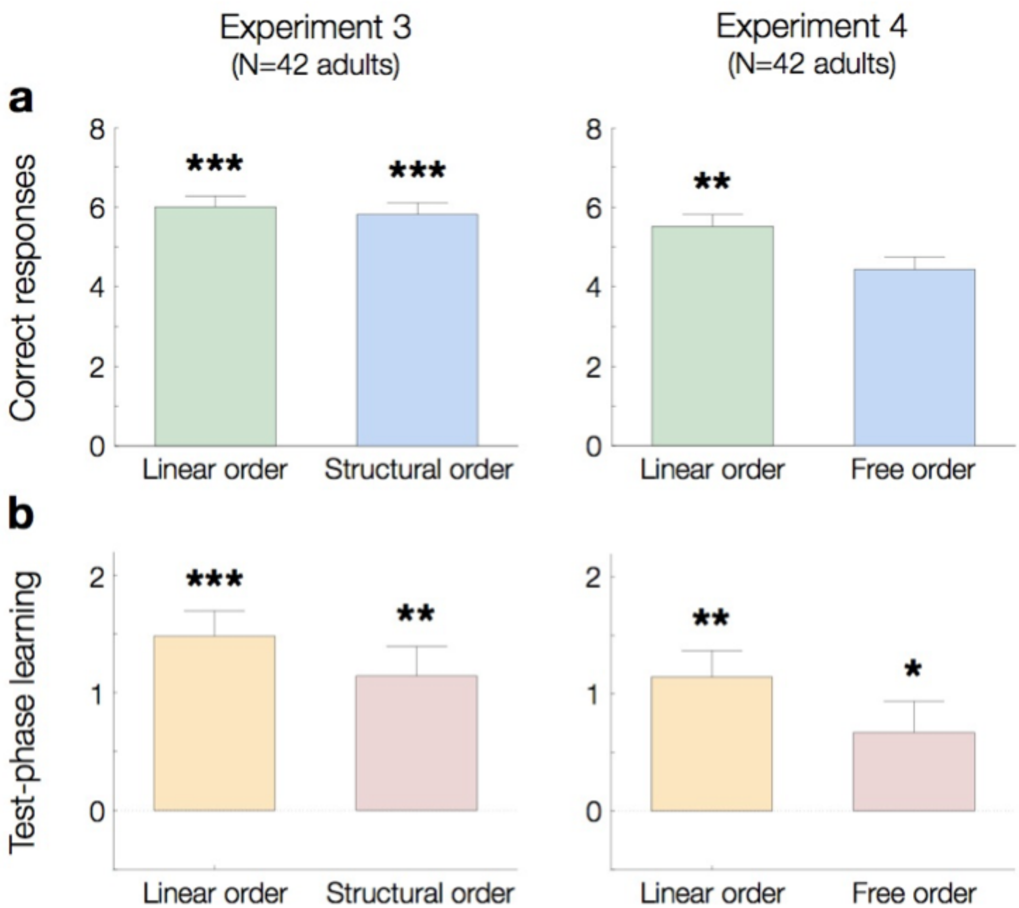


Figure 2.5: Results of Experiments 3 and 4. Panel (a) presents the average number of correct responses, and panel (b) presents test-phase learning effects. Error bars denote Standard Error. Significance codes: p< 0.05*; p< 0.01**; p< 0.001***. Asterisks indicate statistically significant effects (Table 2.2) in one-sample Wilcoxon signed-rank test relative to chance (correct responses, panel a) and Wilcoxon signed-rank tests comparing the second and the first half of the test phase (learning index, panel b). Significance codes: p< 0.05*; p< 0.01**; p< 0.001***.

LO than FO/SO. Learning was not affected either by the grammar they were exposed to during the training phase, nor by their first language. Our results show that structural order constraints do not operate in the same way, assuming they are at all present or active, in adults and in preschool children. Moreover, overall performance seems to be modulated by adult's L1 (unlike children) but learning in the test phase does not.

		SO/FO Correct responses	LO Correct responses	SO/FO Test-phase learning	LO Test-phase learning
Experiment 3	M (SE) Z p r	5.81 (0.298) 3.7 0 0.57	6 (0.285) 3.81 0 0.59	1.14 (0.25) 3.23 0.001 0.5	1.48 (0.214) 3.84 0 0.59
Experiment 4	M (SE) Z p r	4.43 (0.313) 1.25 0.212 0.19	5.52 (0.29) 3.35 0.001 0.52	0.67 (0.27) 2.24 0.025 0.34	1.14 (0.221) 3.36 0.001 0.52
Italian (Exp. 3)	M (SE) Z p r	5.54 (0.366) 2.7 0.007 0.58	5.73 (0.45) 2.57 0.01 0.55	1.36 (0.34) 2.55 0.011 0.54	1.63 (0.28) 2.73 0.006 0.62
German (Exp. 3)	M (SE) Z p r	6.1 (0.48) 2.62 0.009 0.58	6.3 (0.33) 2.83 0.005 0.63	0.9 (0.37) 1.98 0.047 0.44	1.3(0.26) 2.75 0.006 0.62
Italian (Exp. 4)	M (SE) Z p r	3.64 (0.34) 1.1 0.271 0.23	5.18 (0.46) 2 0.046 0.43	0.63 (0.388) 1.55 0.121 0.33	1.18 (0.325) 2.42 0.016 0.51
Polish (Exp. 4)	M (SE) Z p r	5.3 (0.4) 2.36 0.018 0.53	5.9 (0.31) 2.7 0.007 0.6	0.7 (0.39) 1.61 0.107 0.36	1.1 (0.32) 2.41 0.016 0.54

Table 2.2: Mean (M), standard error (SE) and effect size (r). For SO/FO order and LO the statistics are one-sample Wilcoxon signed-rank tests relative to chance level (4), for the learning statistics are Wilcoxon signed-rank tests comparing the first and second half of the test phase.

2.4 General discussion

Words may occupy different positions within a sentence depending on how they are syntactically related to other constituents. This principle, known as structural order, is a key universal property of natural languages: no language obeys a linear order principle whereby constituents have fixed positions within strings. Does structural

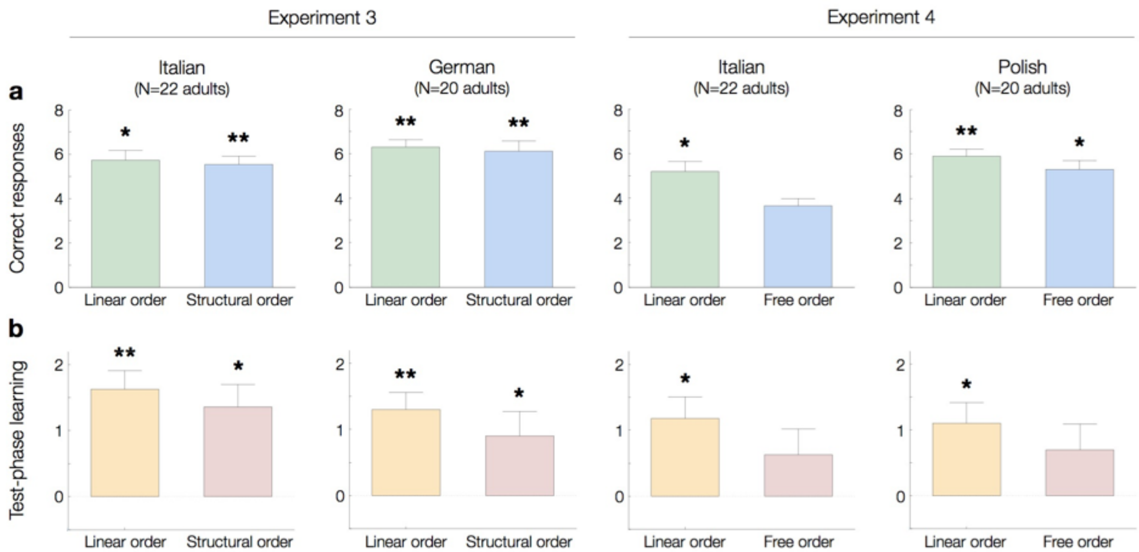


Figure 2.6: Results of Experiment 3 and Experiment 4 for Italian 1 (SO/LO1), German (SO/LO1), Italian 2 (FO/LO2) and Polish (FO/LO2) language groups. Error bars denote standard errors. Asterisks indicate statistically significant effects (Table 2.2) in one-sample Wilcoxon signed-rank tests relative to chance (correct responses, panel a) and Wilcoxon signed-rank tests comparing the second and the first half of the test phase (learning index, panel b). Significance codes: $p < 0.05^{*}$; $p < 0.01^{**}$; $p < 0.001^{***}$.

order constrain language learning, and if so, how? Our results suggest that children are not unconstrained learners. They showed a preference for the grammars they were exposed to during training, and significantly more so for structural order and free order grammars, as compared to linear order grammars. This points to the existence of cognitive constraints or biases on learning that either favors the typologically plausible pattern of constituent order, or disfavors the implausible patterns.

It can be argued that the structural rule used in our experiment is compatible with a hierarchical string organization, but does not require it: it is sufficient for learners to understand that the presence of function words is conditioned by the presence of content words, and treat that as strictly local statistical rule. Note however that LO grammars can also be analysed as relying on local rules, e.g., between the first F and the first C words, and between the last F and the last C words, or in the FCF pattern at the beginning of each string in the first linear order grammar. Also, the free order grammar has no clear local alternation pattern, but seems still easier to learn than a linear order grammar.

Human languages differ in amount of word order flexibility they allow. Some languages, like Italian, have a relatively rigid word order, i.e., Italian uses word order to encode grammatical relations between words. Other languages like German and Polish encode grammatical relations not only through word order but also by means of case marking. Languages with rich case-marking systems typically have a more flexible word order. We tested native speakers of three typologically different languages (i.e., Italian, German, and Polish) with varying degrees of flexibility in word order and different emphasis on morphology, and we found support for the notion that learning constraints are only partly dependent on the first language of learners.

More specifically, we observed that performance in discriminating grammatical from ungrammatical strings was not affected by L1 (as is indicated by a lack of main effects of, or interactions with, L1 in the ANOVA), while learning during the test phase did differ between language groups. The latter finding could be attributed to differences in how linear order structures are learned across languages. One possibility here is that two sets of constraints coexist, determining the learnability of structural and linear order systems, respectively. The former constraints may be invariant across language types, but the latter may be rather more open to L1 influences. Performance of Italian children differed compared to Polish and Austrian children, and we hypothesize that this may be due to that fact that linguistic abilities are acquired in different developmental phases for different languages, and that the timing and duration of their sensitive periods may differ (Friederici, 2005). Another explanation could be related to the fact the utterances in Italian often start with function words. Gervain et al. (2008) found that the position of frequent and infrequent words in a language is strongly related with the head-complement order and children already as young as 7 months old have the ability to compute the frequencies to set the parameter. It suggests that Italian children prefer frequent syllable to occue at the beginning of the units in Italian, especially in utterance-initial positions. Rohrmeier et al. (2012) and Gervain et al. (2013) showed that language experience affect performance of learners in AGL experiments. To what extent language experience plays a role in language processing remains an open question. The design of our experiments does not all us to conclude why we observed the difference in performance in the test phase between the three language groups. However they suggest that cross-linguistic research may be important for determining which constraints are shared across languages, and which are instead language-specific.

Our results suggest that the learning constraints of adults seem to differ from those of children. For adults, discrimination performance was affected both by the grammar they were exposed to and by their L1, while learning during the test phase was not affected by either of these two factors. In contrast, children's performance was modulated by the grammar they were exposed to during the training phase, though not by their L1. Importantly, only learning during the test phase in children was affected by both grammar and L1. Adults responded to the feedback during the test phase by improving their performance on LO grammars, whereas children found SO/FO grammars easier to learn with feedback, showing a dis-preference for LO grammars by 'unlearning' or disengaging with them (Experiment 4). This pattern of results may have implications for understanding why children are more efficient at learning new languages whereas adults struggle in the same task.

It has been proposed that children and adults rely on different processes and implicit strategies for abstract rule extraction, as well as for building linguistic competence in general (Newport, 1990; Hudson Kam & Newport, 2005; Hudson Kam & Newport, 2009). Traditionally one could see our results pointing to a putative 'critical period', after which the ability to acquire language changes. Its existence however has been highly contested suggesting that adults rely on different learning mechanisms (Birdsong, 1992; Perani et al., 1996). Recent EEG studies show that second language learning can reach the same level of fluency and control as L1 (Friederici et al., 2002; Morgan-Short et al., 2012). Already Piaget (1926) suggested that the difference found between adult and children language learning could be a result of maturation of the brain together with the fact that for children language learning is a part of general task trying to understand the surrounding world. This could partially explain the different facility of language learning.

Our data in adults' experiments can also be explained by the interplay of several factors: prior linguistic knowledge, cognitive control, over-learning, and domain-general cognitive abilities interfering with learning certain aspects of language. It is quite possible that the same structural order constraints are also available to adults, but their function is overridden by parallel constraints governing learning in other domains, e.g., linear order in arithmetic. Yet another possibility is that adults, when learning a new language, are selectively suppressing their knowledge of L1, especially if they are explicitly made aware that they are learning a new language. The differences found in our experiments between children and adults are unlike from what is often found in other AGL experiments (Hudson Kam & Newport, 2009). Several studies have found that learning constraints in adults pattern with typological asymmetries. Musso et al. (2003) reported Broca's area activation when learners processed rules of 'possible' languages relative to 'impossible' languages. However, as in our study, there was no difference in behavioral performance between the two types of language. Adults in our study were tested using the same procedure and stimuli as children. It is possible that the number or the length of the strings we used was not sufficient for adults to show learning differences: their overall better performance might result from overcoming learning constraints with simply domain-general mechanisms. Differences could also come from the increased memory load in our experiments as we used relatively big lexicon. Frank & Gibson (2011) suggested that memory constraints can play a role in AGL learning. Another possible explanation for the difference in performance between the two groups is the time it takes to learn new linguistic rules. It is possible that adults need less time to require the grammatical rules and do not use the test phase to further improve their new knowledge, while children use the test phase and its feedback to further learn the rule. Further work is needed to clarify the nature of the observed performance differences between children and adults.

In our experiments we used a more natural AGL paradigm by introducing a key property of natural languages: the distinction between function and content words. After a mere 5 min of exposure to artificial languages, both children and adults inferred, albeit with errors, the order of constituents in strings, using only the frequency and length of pseudo-words as cues. Our results mirror earlier studies suggesting that the position of function and content words is used by learners to extract the constituent order of input strings (Braine, 1966; Morgan et al., 1987; Valian & Coulson, 1988), and to generalize it to novel strings. This provides further evidence for the view that word frequency and the distribution of function words may boost the learnability of languages by signaling the structural position of syntactic units (Hochmann et al., 2010; Gervain et al., 2013).

The main goal of our experiments was to assess the role of learning constraints favoring structural order, or disfavoring linear order, though not to determine their nature or origin. Here, we have provided evidence for the existence of these constraints in preschool children, and for their role during artificial language learning. One point is that rules we used to design our grammars introduce a clear distinction between the two, however we do know know for certain what rules children extracted from the stimuli. It is possible, even though we tried to eliminate potential confounds, that the rules they extracted were based on some other regularities, we cannot exclude that possibility. Culbertson et al. (2012), among others, claimed that, if a constraint or bias exerts pressure on learners to acquire structures that follow certain patterns, the grammars in use would keep changing (or would be selected) to satisfy that constraint. As a result, the performance of learners and cross-linguistic patterns will mirror each other. Here, we presented data on the existence of constraints in favor of structural order, or against linear order, as found in every natural language. This highlights a connection between language learning, typology, and language change (Kirby et al., 2004; Yang, 2004; Wilson, 2006).

The fact that we found effects of learning biases in individual participants in a single experimental session seems inconsistent with theories suggesting that typological differences stem from factors external to cognition, such as historical, geographical and cultural determinants (Evans & Levinson, 2009; Levinson & Evans, 2010; Dunn et al., 2011; Pagel et al., 2013). These proposals would presumably explain the sort of asymmetries in artificial grammar learning observed in our experiments as being a consequence of the fact that children, as well as adults, already possess a native linguistic competence which eventually plays out in the recognition task. Minimally, our results suggest that language learning is not unconstrained. A major challenge for future research is to understand these learning constraints in greater detail, focusing on their origin and scope. This would then allow us to describe how exactly learning constraints interact with maturational, cultural and historical processes to shape natural languages.

Chapter 3

Horizontal and vertical transmission of compositional languages via iterated signaling games

Language change may provide key insights into questions about the nature of language, in particular the role of cognitive and social processes in shaping syntactic and semantic systems. Here, we introduce a novel experimental paradigm – the iterated signaling game (ISGs) – designed to investigate how simple artificial languages ('codes') emerge and change in the course of horizontal coordination between individuals and of vertical transmission across generations. First, we provide a proof of concept that ISGs are a working laboratory model for language change. We show that players consistently converge on a common code after repeated signaling rounds and that the shared code is effectively transmitted and partly modified across generations. Second, we establish a baseline of results for further research using ISGs. We find that compositionality can be imposed on codes by the first generation of players and is maintained by subsequent generations. However, other variables, in particular the degree of coordination among players and how faithfully common codes are transmitted, show a cumulative increase across generations. Our study adds to ample evidence that certain linguistic universals can be studied in the laboratory using models of code transmission in diffusion chains.

3.1 Introduction

In recent years a novel framework has emerged in which languages are regarded as dynamic systems that change across individuals and over time (Beckner et al., 2009; Hruschka et al., 2009). Within this framework, issues arise as to how exactly this variation originates, and what constrains language change. It is likely that principles of human cognition shape, albeit only weakly and indirectly, the evolution of language (Chater & Christiansen, 2010). However, it remains an open question at which stage of cultural transmission such constraints play out: is it primarily during language learning (i.e., during the *vertical transmission* of language from parents to offspring), or is it during language use among peers (i.e., during *horizontal transmission*), or a mixture of both? Some changes in linguistic systems are likely to arise from the individual processes of learning and development (Niyogi 2006), but language is also used in communicative interactions, and therefore its structure may at some level reflects aspects of our social life (Tomasello, 2008; Croft, 2009).

The aim of this study is to introduce a new laboratory model for studying the extent to which properties of natural languages, such as compositionality, are shaped by coordination and communication processes, and how linguistic structure changes through vertical and horizontal transmission. This laboratory model is the *signaling game* (Lewis 1969, Skyrms 2010, Moreno & Baggio 2014). In our experiments, we constructed parallel diffusion chains with several participants ('generations') playing signaling games. In each game, players agree on a *common code*, i.e., a simple artificial language where 'utterances' (the signals) denote events. The common code is used by players in the next game (the next 'generation') as a basis for a new agreement. This iterative procedure allows us to track the emergence and the evolution of referential codes across generations. Although the codes studied here are much simpler and less flexible than natural languages, they do share some key features, such as the presence of expressions denoting features of objects (shape and color) and of events (motion), and a certain degree of structural organisation (compositionality) of signals. In what follows we discuss language change in more detail which is the main theme underlying our research. We introduce our methods by first briefly reviewing previous research on code transmission using iterated learning, and by next presenting signaling games as a novel laboratory model for the study of language evolution.

3.1.1 Language change

In order to model language change, one must consider several interacting constraints, as languages change at nearly every level of organization (Beckner et al., 2009), and may partly do so to meet the demands of processing and communication (Pagel et al., 2007). In some theories, language change is seen as the result of interactions between individual learning and biological evolution (Pinker & Bloom, 1990). More recently, cultural evolution and social interaction have been regarded as factors that may also affect language change, effectively reviving aspects of traditional historical linguistics (Greenberg 1959, Ringe & Eska 2013). In some of the new models (Kirby 2000), what characterises language transmission is that the input which speakers are receiving as their learning material is simultaneously the output produced by other speakers who underwent the same acquisition process. Other models (Hruschka et al., 2009) have emphasized the role of coordination and communication as *goal-oriented processes*. Languages are in constant flux, reflecting the changing demands and characteristics of the communities that use them. Can communicative goals explain which aspects of language change? How does coordination impact on the acquisition and transmission of languages? Are language universals, such as compositionality, partly motivated by social interaction and communicative constraints?

Pagel et al. (2007) observed that higher frequency words are more stable over time than lower frequency words. Words that are more important for communication purposes are those that evolve more slowly. Recent findings by Futrell et al. (2015), based on an analysis of parsed corpus data from 37 languages, supports the view that certain aspects of syntax may have evolved to minimize processing effort and to make communication as efficient as possible. This line of work has provided evidence that languages are affected by communicative pressures, and are to some extent shaped to satisfy communicative needs. Fedzechkina et al. (2012) showed that learners of an artificial language restructured their inputs to increase the communicative efficiency of the target systems within a single experimental session. That result established the important principle that learners alter the inputs they are exposed to if that allows for greater processing ease. Focusing on the emergence of linguistic structures, Kirby et al. (2008) showed how compositionality arises from repeated cycles, where learners acquire and change a language based on the productions of the previous generation of learners. This line of work provided several important insights for experimental and computational research. Simulated communication games have shown that pairwise interactions within a population of agents can prompt the emergence of a shared set of form-meaning associations (Puglisi et al., 2008). Clark et al. (2008) have provided evidence that constraints on word order follow from simple models of learning and social interactions. Moreover, constraints on the processing can give rise to regularities as a result of transmission across generations (Reali & Griffiths, 2009), thus confirming previous historical findings (Bybee, 2007). Although these studies emphasize limited aspects of the complex interplay between social and cognitive factors during language change, they serve as a starting point for providing a more complete picture of how specific assumptions on language learning, processing and use may shape our understanding of language structure.

Viable laboratory and computational models of language change rest on some assumptions. Firstly, for a linguistic form to spread, it needs to be *learnable*. Secondly, learners must exhibit a *bias* (i.e., a preference), however weak, in favor of that form. Such biases may derive from, among others, processing (some forms are less costly to generate or parse) or use (some forms result in more efficient communication) (Kirby et al., 2008; Kirby et al. 2015). Thirdly, languages may change as a result of *conflicting pressures* on speakers and hearers. More specifically, speakers may adopt a principle of *production economy*, and hearers may opt for *explicitness and clarity* (Cooper, 1999; Christiansen & Chater, 2008; Fay & Ellison, 2013). Languages therefore change during transmission, adapting to the cognitive constraints and biases that may play out in the processes of language learning and use by speakers and hearers. Moreover, humans have evolved flexible learning algorithms that track rapid changes in target languages (Baronchelli et al., 2012), despite the fact that words that are replaced more rapidly are those that are used less frequently (Pagel et al., 2007; Pagel et al., 2013), and that higher rates of change are associated with language splitting (Atkinson et al., 2008). It has moreover been suggested that cultural evolution is a factor that determines word order and related aspects of morphosyntax (Dunn et al., 2011). Furthermore, patterns of stability and variation of certain linguistic structures seem to accord with processes of cultural transmission (Dediu & Levinson, 2012).

Social learning is a key ability underpinning cultural transmission (Tomasello, 1999; Richerson & Boyd, 2005). Humans have developed a unique capacity to learn from others that seems crucial for human ecological success (Boyd et al., 2011; Rendell et al., 2011). Communication systems have been found in many animals, but human languages stand out for their complexity and diversity, and for the variety of aims they can serve (Wilson & Sperber, 2002). In studying language change as a result of social learning and transmission between individuals, the focus becomes the progressive modification of information provided to one generation by previous generations. To investigate how this process unfolds, and how cognitive and social constraints may shape languages, leading to the emergence of universal trends, a laboratory model called Iterated Learning was introduced.

3.1.2 Iterated learning

Iterated Learning (IL) was proposed as a model of the processes underlying language evolution (Kirby & Hurford, 2002) based on experimental and computational results showing that IL can lead to the emergence of compositional structure. IL is a process whereby an individual acquires a behavior by observing a similar behavior in another individual who acquired it in the same way. The aim of the IL approach is to simulate the phenomena of interest on a small scale, allowing researchers to investigate how, for example, languages change over generations as a result of repeated transmission. One of the most popular paradigms in IL uses diffusion chains, originally developed by Bartlett (1932). The first person in a chain is given some information and attempts to pass on this information to the next person in the chain. The output is then relayed to the second person who undergoes the same procedure, and so on across the chain. The purpose is to determine to what extent original material is faithfully transmitted. Bartlett found that along the chain the information became degraded and corrupted. Mesoudi and Whiten (2008) argue for the superiority of transmission chains over the traditional single-generation memory experiments as a method for studying cultural transmission. IL has been applied in various computational simulations (Kirby, 2001; Smith et al., 2003; Smith, 2004; Reali & Griffiths, 2009; Swarup & Gasser, 2009; Perfors & Navarro, 2014) and has been used to study instances of cultural evolution such as birdsong (Feher et al., 2009), artifacts (Caldwell & Millen, 2008), stereotypes (Kashima, 2000; Martin et al., 2014) and language (Kirby et al., 2008). IL may provide insights into learning biases that cannot be observed in experiments with individual subjects: biases that are weak in individuals are amplified in diffusion chains (Kalish et al., 2007; Kirby et al., 2007; Griffiths et al., 2008; Smith & Wonnacott, 2010). In the field of linguistics, IL is often combined with artificial language learning to investigate how communication systems and languages may evolve (Scott-Phillips & Kirby, 2010; Galantucci & Garrod, 2011; Galantucci et al., 2012).

A major difference between IL and our model of language transmission – iterated signaling games (ISGs, see below for further details) – is that transmission in ISGs is always based on interaction between senders and receivers. Modeling studies have suggested that interaction and feedback are critical for language change (Steels, 2003). One of the first applications of diffusion chains to investigate the balance and interplay between expressivity and learnability in language evolution was Kirby et al. (2008). In that study, the semantic space consisted of 27 combinations of object color, shape and motion. The first participant learned a set of word-object pairings, and had to generalize it to novel objects by producing appropriate labels. A subset of the labels produced by the first subject was then presented to the second participant as training material, who acquired the mappings and applied them to a novel subset of objects. Subjects never interacted and were not aware that they were part of a diffusion chain. The languages that emerged after several generations were learnable but ambiguous: some signals were used to denote different objects. This result was replicated in other studies (Perfors & Navarro, 2014; Kirby et al., 2015). In a second experiment (Kirby et al., 2008), the ambiguity was artificially filtered out from the training data, and only then the emerging languages were compositional. This artificial filter was introduced as a substitute for communication. However, more recent research by the same group introduces coordination and communication directly into the IL scheme (Kirby et al., 2015). Two participants are trained separately on an input language by observing the mappings between words and their referents. Afterwards, they take turns as speaker and listener in a series of interactive communication rounds. The output language is used as the training set for the next generation. Communication is introduced *within* each generation. Still, there is no interaction-based transmission *across* generations. Unlike in the training of Kirby et al. (2015), in fact language acquisition takes place by imitating, observing and interacting with others. Communication between individuals interacting to reach a certain goal using IL was also used in studies of the structure of graphic communication systems (Healey et al., 2002, 2007; Garrod et al., 2007; 2010; Theisen-White et al., 2011). Despite important analogies between graphic and verbal communication, graphic symbols lack combinatoriality, compositionality, and duality of patterning (Hockett, 1960). Moreover, in natural language, the relation between an expression and its meaning is typically (and universally) arbitrary, and not iconic.

In *seeding* the initial behavior or language is decided by the experimenter, and is then spread and modified by participants within the diffusion chain (Horner et al., 2006; Whiten et al., 2007; Flynn & Whiten, 2010; Whiten & Flynn, 2010; Nielsen et al., 2012). Diffusion studies primarily aim at understanding how well the seeded patterns spread. In studies with adult participants, for whom cultural transmission is already an established behavior, the issue is how and why behavioral patterns change over time (Mesoudi & Whiten, 2008). Studies of language transmission typically seed with random strings that denote 'holistically' (i.e., as fixed labels). This choice stems from theories of 'proto-language' as a holistic signaling system (Wray, 1998; Kirby, 2000; Arbib, 2005; Fitch, 2007). These theories state that proto-languages initially had only discrete proto-words representing concepts, and forming at most short, unstructured proto-word strings. This notion, however, is problematic. As Bickerton (2003) points out, it would be difficult for speakers to settle on an agreed meaning when utterances are holistic. Others suggest that early communication evolved from a proto-language closer to actual languages, albeit much simpler, i.e., with a proto-lexicon but no syntax (Bickerton, 1990; 2003; Jackendoff, 2002; Tallerman, 2007). As Bybee (2012) also has observed, grammatical structures are formed by unification of adjacent elements, not by breaking complex elements apart. Therefore, if the aim is to demonstrate language emergence and change through cultural transmission, seeding with holistic languages may not be the most ecologically plausible approach. Indeed, there is no evidence that complex signaling systems originate from a random, complex, unstable state. Seeding with structured systems, on the other hand, has the advantage of steering clear from the question 'what was there initially?': a plausible intermediate evolutionary state of the language is used as a starting point of investigation. That is our approach here.

3.1.3 Signaling games

Experimental research on dialogue has produced a substantial amount of knowledge about the principles underlying language use (Clark, 1996; Brennan & Hanna; 2009). One upshot of this line of research is that, to understand the structure of languages, one needs to focus both on individual cognitive constraints and on social interactions. As suggested by the interactive-alignment model of dialogue by Pickering and Garrod (Garrod & Pickering, 2004; Pickering & Garrod, 2004), convergence, in which one speaker's output is made to increase in similarity to another speaker's output, occurs when a speaker adjusts her behavior to that of her interlocutor's for the purposes of communication (Soliz & Giles, 2014). So far, phonetic convergence has been observed in interactive and non-interactive situations, and morphological conformity has also been observed (Beckner et al., 2015). Information transmission was more effective in interactive compared to non-interactive situations (Tan & Fay, 2011). Collaborative accounts propose that individuals interact to ensure that what is said is also agreed upon and properly understood by both parties. Therefore, meaning is bi-directionally negotiated. Already Clark and Wilkes-Gibbs (1986) found that dyadic conversational references are established over extended exchange. Typically, such exchange consist of cycles of interaction resolved only when there is mutual agreement on referents by both parties. An exchange typically starts with idiosyncratic descriptions, and exploits subsequent interactions to adjust mismatching referents. Similar results have been obtained in studies of the emergence of communication (Galantucci & Garrod, 2011; for review see Galantucci et al., 2012). Finally, specific brain areas that do not entirely overlap with those involved in language processing, are engaged during coordination and communication (Ramnani & Miall, 2004; Carrington & Bailey, 2009; Willems et al., 2010; Enrici et al., 2011; Stolk et al., 2014). These data are broadly consistent with the independent contribution of coordination and communication to shaping linguistic structure.

In IL, participants learn a miniature artificial language and are subsequently tested on their knowledge. Their output from the test phase is used as an input for the next participant. In our own studies, we introduced an interactional model of cultural learning – the *signaling game*. Signaling games (Lewis 1969; Skyrms 2010) are a class of dynamic games of incomplete information. What makes signaling games relevant here is that they are a relatively simple yet highly flexible model of how agents update their beliefs based on observed actions (signals), and of how the interaction changes accordingly. In the simplest case, the sender has access to a state of the world, which is unknown to the receiver. The sender sends a *signal* to inform the receiver about the state. The receiver then chooses an *action* in response to the signal, and if the action is appropriate given the state of the world (e.g., bring an umbrella, if it is raining), the trial is successful for both players. If the sender consistently uses a signal in one state and another signal in another state (i.e., what is called a 'separating equilibrium', for arbitrary numbers of states and signals), the receiver can always choose the relevant action in response to a signal. What sender and receiver do is what gives meaning to signals. Signaling game theory can be used to study information flow in networks of multiple players (Skyrms 2009), capturing a variety of real-world situations. Here, we introduce iterated signaling games (ISGs), that is, signaling games played sequentially in *chains* (open and totally ordered) of senders and receivers, in which the receiver in a game becomes the sender in the next game.

3.1.4 Aims of the present study

The aim of our research is twofold. First, we provide a *proof of concept* that ISGs are a viable laboratory model of language change. We show that (1) in each signaling game, players converge on a common code after repeated signaling rounds with feedback, and (2) codes are effectively transmitted and modified across generations. Second, we establish a *baseline of results* for further experimental research with ISGs. ISGs are a highly flexible paradigm as feedback, the roles of the players, the size of the state and signal spaces, signal frequency, and so on, are all easily manipulated. In this sense, the present study is an exploration of the simplest possible ISG paradigm: i.e., feedback is complete, sender and receiver roles are fixed throughout a game, the size of the state and signals spaces are limited and known to players, and states are equiprobable for the sender (all signals are equiprobable for the receiver). Our focus is on the structure of the code as well as on mappings of signals to referents. The constituents of signals and the communication channel are predefined. We also aim to provide a 'conceptual replication' using a different paradigm (i.e., ISG) of previous results (Theisen-White et al., 2011; Kirby et al., 2015) showing that pressure from learning and communication leads to structured codes. The novelty of our approach lies in using the ISG paradigm, in which communicative goals are explicit, feedback is complete and is provided as common knowledge to players and seeding is based on a non-holistic protolanguage. As part of establishing a baseline for ISGs, we aim to replicate earlier results (Moreno & Baggio, 2014; Lumaca & Baggio, submitted) suggesting that, in signaling games in which roles are fixed (i.e., a player is either sender or receiver throughout a game), horizontal code transmission occurs from sender to receivers.

In studies of cultural evolution, two 'modes of transmission' are considered: vertical and horizontal. Here we define horizontal transmission as within-pairs transmission and vertical transmission as across-pairs transmission. Vertical transmission assumes that cultural information flows as in IL *from generation to generation* (Kirby et al., 2008). Horizontal transmission is the emergence of communication systems through interaction *within a generation* of interlocutors: it can be unidirectional (e.g., from sender to receiver) or bi-directional (Galantucci, 2005; Healey et al., 2007; Selten & Warglien, 2007; Scott-Phillips et al., 2009). Both can lead to the emergence of compositionality. Garrod et al. (2010) compared two types of transmission using a version of a graphical communication task. However, as mentioned earlier, their results are not necessarily relevant for language. In ISGs, both modes of transmission are present. This allows us to investigate whether, even with coordination at the horizontal level, the final codes after vertical transmission are more learnable and structured.

One hypothesis that we will be able to test is whether rapid changes in vocabulary occur during horizontal transmission, whereas vertical transmission explains gradual changes in linguistic structure. As suggested by Bickerton (1990), repeated pairwise interaction can lead to the emergence of a compositional structure and linear order already at the outset of the interaction, due to communicative pressure. This does not mean that vertical transmission has little or no influence on structure. However, we expect its role to be less dominant than interaction. The participants' main goal is to communicate effectively. Therefore, our prediction is that certain essential properties of languages, without which communication is difficult or impossible, would emerge quickly within one or a few generations. Other properties, that may refine codes and render them 'optimal' in some respect, will instead appear gradually and change over generations. A key question here is whether compositionality belongs to the former or to the latter set of properties.

3.2 Experiment 1: Compositional semantics

We based our stimuli on Kirby's et al. (2008) experiments. Our paradigm rests upon the assumption that to investigate language change one needs an interactive task. Our participants' explicit goal was to coordinate and communicate successfully. Moreover, they were aware they were part of a transmission chain. Thus, the task in ISGs is not just a memory recall task.

3.2.1 Methods

3.2.1.1 Participants

Thirty eight (38) Italian native speakers (mean age 23.5, age range 19-33, 24 female) participated in Experiment 1. They all had normal or corrected-to-normal vision, and all were trichromats. They were recruited via an institutional website and were monetarily compensated for their participation. Sample size was based on previous research on IL (Kirby et al., 2008). Upon their arrival to the lab, the participants were informed they would play a game with a partner. Participants who played together in a game did not know each other beforehand. Unlike in previous experiments using diffusion chains, participants were aware that the study involved interaction between partners. At the end of the experiment, participants were debriefed about the goal of the study. Participants were organized in 4 transmission chains of 9 generations each. Two participants failed to show up for the sessions, thus the original plan of using 10 generations in each chain had to be abandoned: that is, 36 participants were included in the final data set. The study was approved by the Ethics Committee of SISSA.

3.2.1.2 Apparatus

Each pair of participants was seated in the experimental room at a large desk facing each other, so that each player could not see the other during the session. Each player had their own workplace consisting of a computer screen and keyboard. The screens were aligned back to back, rendering it impossible for each player to see their partner and their screen. Sender and receiver roles were assigned at the beginning of the each session, and were fixed throughout the session. At the end of each session (generation n), the receiver would become the sender in the next session (generation n+1). Thus, chains were constructed. In between two sessions, the receiver was transferred to the sender position, and a new participant was let into the room to play as a receiver. The stimuli were delivered from a single computer controlling both workstations using Presentation (Neurobehavioral Systems).

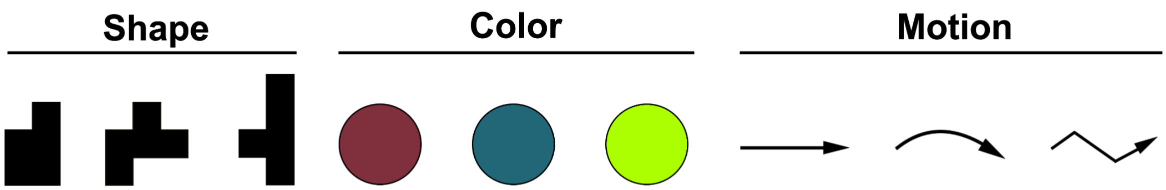


Figure 3.1: Set of visual stimuli used in Experiment 1. Total set consisted of 27 scenes, varied along three dimensions: shape, color and motion.

3.2.1.3 Stimuli

We used the same 3-by-3 stimulus design as in Kirby et al. (2008). The *states* in our signaling games are visual scenes varying in 3 dimensions: shape, color and motion (Fig. 3.1). We used artificial Tetris-like shapes and ambiguous colors to prevent players from transferring (partial) mappings of labels to meanings (e.g., ‘re’ to red) from their native language, which would act as undesirable ‘focal points’ (preplay equilibria) in a game (see Moreno & Baggio (2014) for a discussion). All object shapes had the same number of constituent squares (i.e., 5). All motion trajectories (straight, curved and zig-zaggy) of objects started and ended at the same two locations on the screen. The constituents of *signals* were monosyllables. To prevent players from typing-in labels, and possibly using words of their own native language, the labels had to be chosen by players (details below) from a set of 9 labels shown on the screen. The task consisted of matching scenes resulting from the factorial combination of three feature sets, with three features each (shape, color, motion; 27 possible), to 3-syllable strings as *signals*. For the *sender*, the task consisted of mapping states (visual scenes) to signals, and for the receiver, the task was the reverse (see below). We used closed sets for signals and states, and that enabled us to measure some language-like properties of the emerging signaling systems, i.e., compositionality, given a linear constituent order.

3.2.1.4 Procedure

Participants were organized into 4 vertical diffusion chains with 9 players each, and played a signaling game with fixed roles and multiple signaling rounds. Each player took part in 2 games or sessions of the experiment. In the first session, player A was the sender and player B was the receiver in generation n. Once A and B converged on a shared mapping of signals to states, after a short break a new game started, where player B was now the sender, and player C was introduced as a new receiver, thus establishing generation n+1. Players were not allowed to communicate verbally or otherwise. The experimenter was always present in the room. We seeded chains as follows: the first sender in each chain (generation 1) was trained using a randomly generated bijective mapping of syllables to object features. In each learning trial, the participant was presented with an object, and had to compose a single 3-syllable label describing that object. Feedback was provided, as to which labels and object features were correctly or incorrectly matched in that trial. The training phase finished when the participant learned all the associations. There was no time limit for this task.

Each signaling trial unfolds as follows (Fig. 3.2). The sender is privately shown an event (1s clip) drawn randomly from 27 possible combinations of shape, color and motion, and is asked to send a signal to the receiver denoting the event. To do so, the sender chooses 3 labels from the learned vocabulary: all 9 labels are shown at the same time on the screen in a 3-by-3 grid (duration: self-paced), in a random order in each trial. The sender composes a signal, which is immediately sent to the receiver. Both players simultaneously see the signal the sender has composed. Based on this signal, the receiver has to respond by composing the event he believes the sender has seen, choosing shape, color and motion features in a 3-by-3 grid (duration: self paced) in a different randomly generated order in each trial. Feedback (2s) is presented to both players indicating whether the elements that the receiver has chosen match the elements of the event that the sender has seen.

Studies of coordination games support the idea that feedback is important for the emergence of successful communication systems (Healey et al., 2007; De Ruiter et al., 2010). The game ends when participants reach 60 correct trials, with no constraint on the number of correct consecutive trials. A trial is correct if and only if all three features selected by the receiver in response to the signal match the features of the event as seen by the sender. Therefore, a trial is successful if the message sent by the sender is fully understood by the receiver. Note that the events presented to senders are all equiprobable. This is desirable given the purposes of the present experiment, since in languages more frequent words tend to exhibit less variation (Pagel et al., 2007). Each session lasted approximately 1 hour.

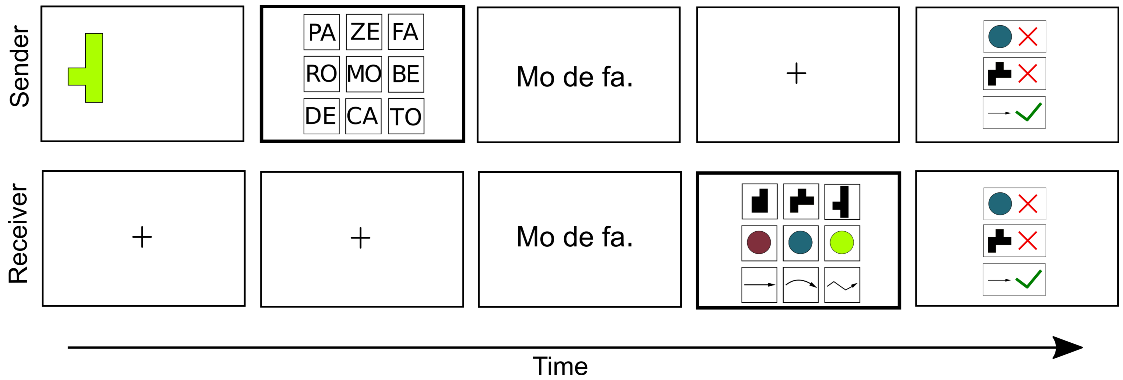


Figure 3.2: An example of trial in a signaling game The top and bottom rows show what the sender and receiver see, respectively. Time flows from left to right. In each trial, the sender sees a scene (clip) and composes a signal (maximum 3 syllables) to communicate the identity of the scene to the receiver. The receiver sees the signal and chooses from the 9-object array the elements that he believes were seen by the receiver. Next, the feedback is presented to players, indicating how accurate the receiver's choices have been. Players use the numerical keypad on the full size keyboard to produce their responses. Bold frames indicate the points at which sender and receiver respond.

3.2.1.5 Data analysis

Our aim is to describe changes in code structure, mappings and regularities during transmission, both within generations and within chains. In particular, we tested for changes in the quality of code transmission and coordination, the gradual emergence of compositionality, and innovation and fidelity in code structure. These measures are described below. The expectation from earlier research is that both transmission and compositionality gradually increase over generations (Kirby et al., 2008). However, given coordination pressure and communicative goals, a structured signaling system may emerge earlier or suddenly.

We used 5 measures: *structure, coordination, transmission, innovation* and *asymmetry*, described in detail below (see Moreno & Baggio, 2014). Following Kirby et al. (2008), we measured code similarity using mean normalized Levenshtein distances between the codes agreed on by one generation, and the codes formed by the next generation, averaging over all meanings. The normalized Levenshtein distance is defined as the smallest number of character insertions, replacements or deletions that are required to transform a character string into another, divided by the length of longest string. The output is a number between 0 and 1, with 1 indicating fully matching codes, and 0 fully mismatching codes.

First, *structure* change was measured using RegMap (Tamariz & Smith, 2008; Tamariz et al., 2010), a measure of the degree of confidence that a signal element consistently predicts a meaning element. A partial RegMap estimates how reliable an association between meanings and signals is. The full RegMap tests whether these mappings are bi-unique, returning a single value for the entire language. Higher scores suggest that similar meanings are consistently predicted by similar signals, as observed in natural language. We expected structure to increase over generations.

Second, *coordination* was measured using the mean normalized Levenshtein distance between codes of corresponding states as adopted by sender and receiver in a game. Values range from 0 (no shared code) to 1 (common code). The coordination measure indirectly reveals how learnable a code may be, and to what extent an agreement on a common code could be found in a signaling game. An increasing coordination along the transmission chain would suggest that codes become more learnable and more easily agreed upon by players.

Third, *transmission* measures the extent to which the codes used by two subsequent generations match. As predicted by earlier research (Kirby et al., 2008), transmission may increase along a diffusion chain if players restructure the code to render it more learnable and less prone to errors. However, if a sender wants to reproduce a code acquired in the first session (playing as receiver), the sender may be faced with the situation where, for example, the receiver does not learn the code. In such a situation, a strategy may be to remap the associations. The relevant measure here is the fourth index – *innovation* – designed to capture how much the code produced by the sender in generation n differs from the output code of the sender in generation n+1. If this value *decreases* over generations, the code is likely to become more learnable.

Finally, we measured *asymmetry* between senders and receivers. McAvoy & Hauert (2015) have demonstrated that interactions between players are often asymmetric. Moreno & Baggio (2014) have shown that, with fixed-role signaling games (in which a player is either sender or receiver throughout a game), coordination labor is divided. Receivers adjust their mappings more frequently, and senders tend to maintain their initial mapping until it becomes the common code. Our objective was to examine the division of labor in code transmission, and possibly replicate the findings of Moreno & Baggio (2014) in the context of transmission chains. Asymmetry is defined for a pair in a game as the number of code changes made by the sender minus the number of code changes made by the receiver, divided by the number of code changes made by both players (total code changes). Asymmetry ranges from -1 (changes are made only by the receiver) to 1 (changes are made only by the sender). The summary of our measures is shown in Figure 3.3.

Players are free to use some or all syllables, and in various ways to denote events, so it is in principle possible that ambiguous languages emerge, in which a single syllable or signal has several meanings. We computed the average number of syllables in each code used across features in each dimension (shape, color and motion) to assess code expressivity.

We applied Page's trend test (Page, 1963) to test for cumulative changes in all of the above measures across generations. Page's trend test is used in data sets where there are at least three samples of repeated measures, and where the hypothesis that is being tested a priori predicts a particular direction of effects. We applied Wilcoxon's signed-rank nonparametric tests to pairwise comparisons between generations. We submitted each measure to a mixed-design ANOVA model with Generations (9 levels) as a within-subjects factor.

3.2.2 Results

We found no decrease or increase in compositionality over generations (Page's trend test on RegMap: L=858, p=0.84). From the first generation, in all chains, codes were highly structured (M=0.87, SD=0.01). There was no difference in structure between generation-1 and generation-9 (G_1: M=0.87, SD=0.02, G_9: M=0.84, SD=0.07; Wilcoxon signed-rank test Z=-0.38, p=0.7). The ANOVA showed no main effect of Generation (F(7,21)=0.08, p=0.94). The compositional codes from the first to the last generation were built based on a fixed linear order of constituents: for example, the first syllable denoted the object's shape, the second movement, and the third color. Our concept of structure does not quite match the complexity of natural language. However, it does conform to standard definitions of compositionality. Previous experiments using IL have shown gradual increases in compositionality during vertical transmission (Kirby et al. 2008, Kirby et al., 2015). Our data suggest that, in a paradigm such as ISGs, with communicative pressure and feedback, code compositionality arises imme-

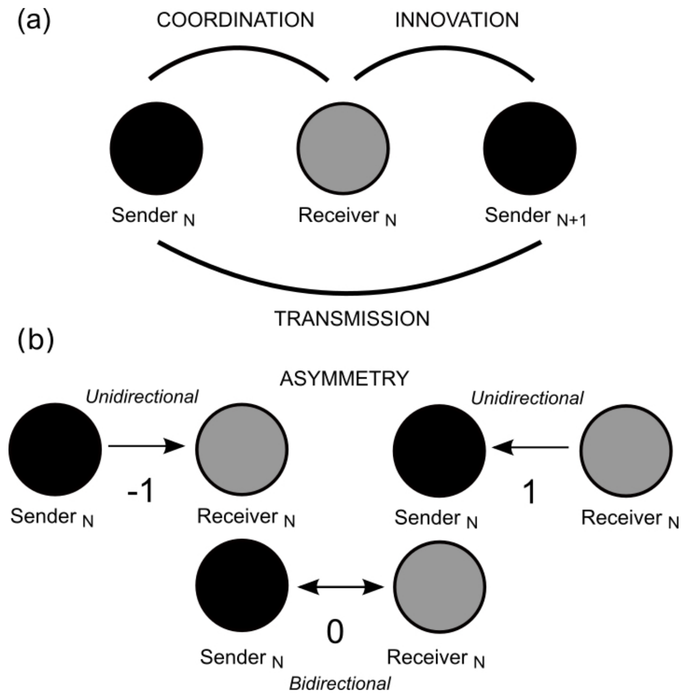


Figure 3.3: The measures used for the analysis.

diately and is maintained across generations.

Figure 3.4 shows how coordination, transmission and innovation change in diffusion chains. *Coordination* increased gradually (L=990, p=0.01), implying that codes were becoming easier for partners to agree upon along a transmission chain. We found an effect of Generation in the ANOVA (F(7,21)=21.76, p=0.019). Coordination was high already in generation-1, indicating that players found it relatively easy to agree on a common code. Coordination in generation-8 (M=0.94, SD=0.1) was only moderately higher than in generation-1 (M=0.79, SD=0.73; Wilcoxon signed-rank test: Z=-1.46, p=0.144). We observed a marginal increase of *transmission* over generations (L=959, p=0.04), suggesting that players in later generations were slightly more accurate in transmitting codes. As with coordination, transmission was moderately higher in the last generation (M=0.92, SD=0.17) compared to the first (M=0.77, SD=0.24; Wilcoxon signed-rank test: Z=-1.069, p=0.28). There was no main ANOVA effect of Generation (F(7,21)=3.33, p=0.16). *Innovation* did not decrease cumulatively (L=662, p=0.312). There was no effect of Generation (F(7,21)=0.313, p=0.62), suggesting that although coordination and transmission were increasing, players continued to change the code throughout the whole chain.

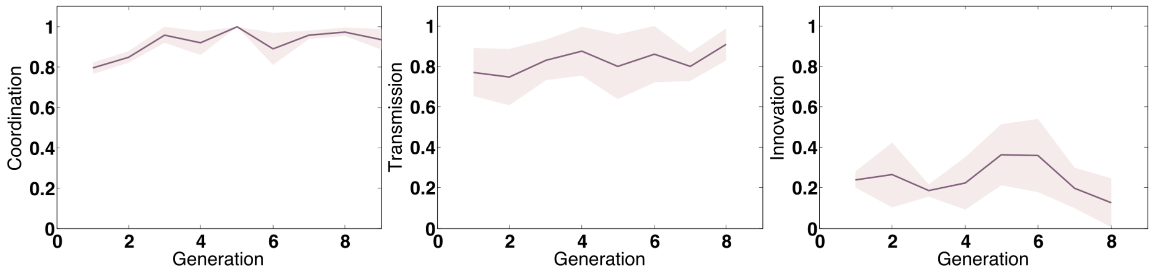


Figure 3.4: The effects of Generation on coordination, transmission and innovation. 95% confidence intervals are shown as shaded areas.

We further explored what kind of changes participants were introducing in the codes. We found that 12% of all changes were due to participants changing the *order of the elements* in codes, 25% of changes arose from players changing codes *across semantic categories* (e.g., exchanging a word associated with color for a word associated with motion), and the greatest number of changes (63%) came from remapping syllable meanings *within a semantic category.* To further investigate the sources of innovation, we looked at asymmetry to understand the division of labor between players during code change. Asymmetry was negative and was significantly different from 0 (M=-0.6, one-sample t-test: $t(35)=35.9$, $p=0.001$), suggesting that receivers tend to adjust their mappings more often than senders.

3.2.3 Interim discussion

Previous iterated learning experiments, in which learnability but not communication exerts a pressure on transmission, showed that ambiguous codes arise in the course of transmission: individual signals are reused to denote more meanings. If, however, ambiguity is removed from training sets, then compositionality arises. This artificial filtering procedure was used in the first versions of the IL paradigm as “an analogue of a pressure to be expressive that would come from communicative need in the case of real language transmission” (Kirby et al. 2008, p. 10684; Kirby et al., 2015). One of our aims was to explore whether introducing communicative pressure *directly* into the experimental paradigm, here signaling games, is sufficient for the emergence of structured codes. Our results are in agreement with Kirby's et al. (2008) hypothesis that it is indeed communicative pressure that leads to structured expressive codes. However, we did not find a gradual increase in structure over generations. Whereas Kirby et al. (2008, 2015) observe convergence to a compositional system, in our study structure emerges already in the first generation, and remains stable over subsequent generations. The difference could stem from the fact that our participants interacted repeatedly with the explicit goal of arriving at a shared system for communication. Our results parallel those of Vogt (2005) where compositional languages emerge rapidly and remain very stable as long as horizontal transmission is present. The signals that emerged in our transmission chains resemble morphologically complex words in agglutinative languages. That is, one syllable denoted the object's shape, one denoted color, and one denoted motion. Interestingly, constituent order in a signal, in which each of three constituents denotes different semantic categories (e.g., the first constituent denotes shape), did not change much after the first sender introduced it. In none of the chains have we noticed a decrease of the total number (27) of syllables.

Earlier research has shown that non-linguistic signs established in subsequent stages of novel communication systems consolidate already established signs (Galantucci, 2005; Garrod et al. 2007; De Ruiter et al., 2010). This occurs also in systems like ours, consisting of only a limited number of signals, suggesting that linguistic structure may emerge in the early stages of the evolution of a communication system. In real life, vertical transmission (i.e., from parents to offspring) is often influenced by learning bottlenecks, whereas horizontal transmission (i.e., between sender and receiver) is shaped by the informational gap between the interlocutors. In order to minimize or close the gap, speakers may rely on a (partly) compositional language from as early as possible in their exchange. The more transparently compositional a language is, the faster alignment occurs. Our data suggest that participants do not exactly reproduce the language: innovation does not decrease over generations. One of the explanations is that, as in the real world, the main goal of a speaker is successful communication, not necessarily the exact reproduction of a target language. Compositionality makes communication more efficient, but faithful reproduction and transmission seems less important so long as the communicative goal is reached.

3.3 Experiment 2: Grammatical functions

In Experiment 1 we established that signaling games are a viable model to investigate code transmission and evolution. We conducted a second experiment to study forms of grammatical change. Specifically, we tested whether the emergence of grammatical categories is partially influenced by communication and learnability pressures.

An established breakthrough in language evolution was the emergence of categories at various levels of linguistic representation (Garrod & Anderson, 1987). Puglisi et al. (2008) demonstrated that agents with basic communication principles may evolve a communication system with linguistic categories. Simple interaction with feedback is sufficient for a self-organized system to emerge that is able to discriminate categories, using only a small set of words. One of the most basic requirements in understanding language is identifying grammatical categories to which words belong. Grammatical categories are classes of words that have a similar syntactic role in sentences. Words belonging to the same grammatical category can be exchanged in a sentence without it becoming ungrammatical.

In the second experiment, we wanted to investigate the emergence of structure and basic grammatical categories. In order to learn the code, participants had to identify words denoting objects ('nouns') or actions ('verbs'), and had to link the grammatical category of the nouns to the semantic categories of agent and patient. In the first experiment, we borrowed the design of Kirby et al. (2008) to see how adding communicative pressure affected the language transmission using shape, color and motion. In the second experiment, our goal was to investigate the emergence of a shared code using two objects and a concrete action that the objects are performing.

Additionally, our aim was to investigate how speakers of different languages (Italian and Polish) restructure the linguistic input. This may allow us to extend the results to speakers of more than one language, and minimize the risk that some learning biases are transferred from speakers' knowledge of their native language. While recent work has revealed universal characteristics in human social interaction (Stivers et al., 2009, Dingemanse et al., 2015), natural languages vary in fundamental ways. They make use of word order, inflection and prosody to transfer meaning, but certain languages rely more heavily on one or another grammatical device for specific functional purposes. Interestingly, Polish and Italian both allow for relatively free word order (Polish word

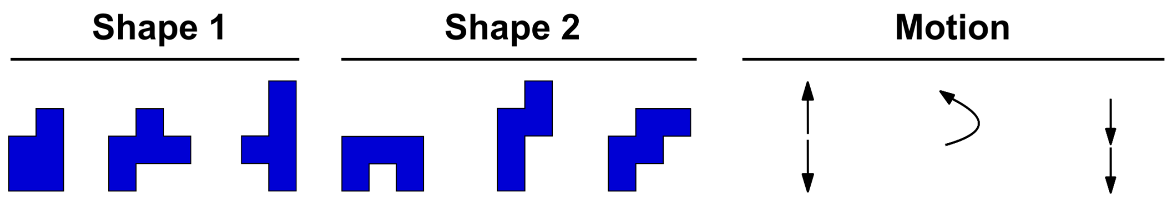


Figure 3.5: Set of visual stimuli used in Experiment 2. Total set consisted of 27 combinations, varying along three dimensions (Object1, Object2, Motion).

order is largely free, while Italian allows for a highly mobile position of the verb) but implement it using different linguistic devices: Polish has rich case-marking system; Italian cases exist only for pronouns.

3.3.1 Methods

3.3.1.1 Participants

Thirty seven Italian speakers (mean age 24.7, age range 19-30, 27 female) and thirty nine Polish speakers (mean age 22.3, age range 18-25, 22 female) participated in Experiment 2. They were recruited via an institutional website and were paid for their participation. As we had non equal number of generations in different chains, to match all the chains in the analysis we included 8 transmission chains of 9 generations each: 4 chains with Italian speakers, and 4 with Polish speakers.

3.3.1.2 Stimuli

The stimuli consisted of 6 shapes (divided into two pools of 3 shapes) and 3 motions. We presented participants with visual scenes varying in three dimensions: one out of three shapes from the first pool of shapes, one out of three shapes from the second pool of shapes, and a motion type (Fig. 3.5). The tetris-like shapes had the same characteristics as in Experiment 1. All motions were designed to start in the same location of the screen. The shapes have been previously screened using a separate sample and the most ambiguous ones were chosen. Signals were constituted by monosyllables with the same properties as in Experiment 1.

3.3.1.3 Apparatus, Procedure and Analysis

The same as in Experiment 1.

3.3.2 Results

Compared to the first experiment, the second was more difficult: players needed on average 27 more trials to agree on a code (Exp1: M=96, SD=14.49; Exp2: M=123, SD=13.54). Six objects may have been more difficult to differentiate, therefore players needed more time to learn the codes. As in Experiment 1, compositional structure emerged already in the first generation (M=0.85, SD=0.019), and remained stable and on a high level throughout the transmission chain (Generation-9: M=0.84, SD=0.011). Compositionality did not increase over generations (Wilcoxon signed-rank test, Z=-1.089, p=0.27; Page's Trend Test on RegMap data: L=608, p=0.96). There was no main effect of Generation in the ANOVA (F(8,56)=0.128, p=0.75).

Similarly to Experiment 1, we found a cumulative increase in coordination (L=994, p=0.013) and transmission (L=726, p=0.01). This finding is not supported by pairwise comparisons between the first and last generation, which show only marginal effects (Coordination: G_1, M=0.66, SD=0.17, G_9, M=0.98, SD=0.01, Wilcoxon signed-rank test Z=-1.826, p=0.068; Transmission: G_1, M=0.54, SD =0.04, G_8, M=1, SD=0, Z=-1.461, p=0.15). We found an effect of Generation in ANOVAs for Coordination (F(8,56)=10.5, p=0.04) and a marginal effect of Generation for Transmission (F(8,56)=7.47, p=0.07). These data show that codes were becoming easier to transmit, and that coordination between partners was also easier. One of the key differences between Experiment 1 and Experiment 2 was the cumulative decrease in innovation in Experiment 2 (Page's Trend Test: L=716, p=0.012). This suggests that codes tended to become more stable and less susceptible to change, also as a result of transmission. The difference in innovation between the first and last generation was only marginally significant (G_1: M=0.27, SD=0.07; G_8: M=0.05, SD=0.05, Wilcoxon signed-rank test Z=-1.86, p=0.06), and there was no main effect of Generation in the ANOVA (F(8,56)=2.19, p=0.23). Asymmetry was negative (M=-0.67, t(71)=11.8, p=0.001). As in Experiment 1, structural asymmetry was unidirectional with the sender imposing the word order, leaving semantic asymmetry to be influenced by both sender and receiver.

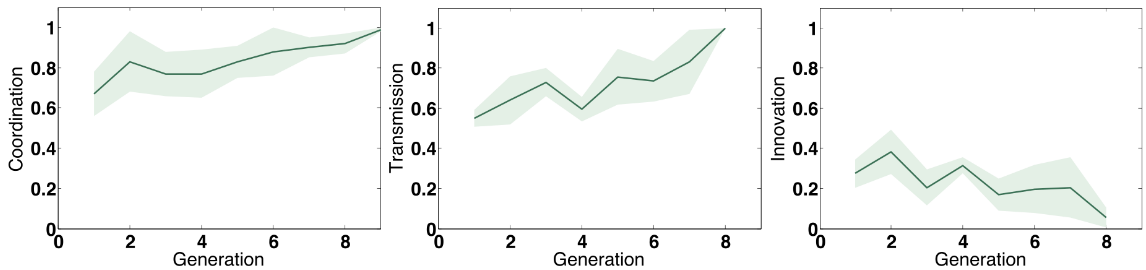


Figure 3.6: The effect of Generation on (a) Coordination (b) Transmission and (c) Innovation. 95% confidence intervals are shown. All three measures changed cumulatively over generations.

We did not find differences between speakers of Polish and Italian. We submitted all relevant variables into ANOVA with the speaker's native language (Polish or Italian) as a between-subjects factor. For none of the variables did we find a main effect of the language of participants (Structure: $F(1,8)=2.33$, $p=0.17$; Coordination: $F(1,8)=2.94$, $p=0.13$; Transmission: $F(1,8)=3.34$, $p=0.09$; Innovation: $F(1,8)=1.86$, $p=0.22$).

In 5 out of 8 chains (3 in Polish, 2 in Italian) players gradually modified the codes so as to use word cues (vowels) to group elements into semantic categories: for example all Motions were mapped to words that included the vowel 'o': 'ro', 'to', 'mo' (see Fig. 3.7). Interestingly, this did not happen in Experiment 1, even though our signals were the same in both experiments. These 'morphological' markings may seem redundant, for linear word order already establishes category-specific positions in a signal: e.g., the syllable associated with motion comes always second. It is possible that, as task difficulty and processing demands increased compared to Experiment 1, linear word order was insufficient to establish a 'safe' communication system. Thus, participants turned to other marking mechanisms resembling those found in natural languages. As in Experiment 1, mappings were stable across semantic categories, but continued to change within each category as players negotiated meanings in dyadic interactions.

3.3.3 Interim discussion

In the second experiment too, we found that codes maintained their expressivity and elements retained distinct labels over generations. The codes became more stable as a result of transmission, as shown by the cumulative increase of innovation: this was a difference between the two experiments. We hypothesize that innovation decreases in the second experiment due to greater task difficulty. Participants turned to other linguistic devices to transform the input codes into more learnable ones. At the same time, while innovation decreased, coordination and transmission increased without affecting code structure. As the data for structure show, communicative goals render codes both expressive and learnable also in Experiment 2.

The most interesting finding of the second experiment is the emergence of what may be called basic morphosyntax. The final codes resembled agglutinative languages: vowels here denoted grammatical categories, whereas consonants denoted the exact associations to items within a category (Fig. 3.7). The systematicity that we found in the mappings between signals and referents is a universal of natural languages, and is also present in the morphological structure of the lexicon (Monaghan et al., 2014). One of the functions of morphology is to highlight grammatical relations. In the first experiment, 3 dimensions (shape, colour and motion) were easily distinguishable and linear order was sufficient to generate meaning. Comrie (1981) proposed that morphology appeared at a relatively late stage of grammar evolution. Heine & Kuteva (2002) suggest that first there was a distinction between verbs and nouns, and only with an increase in the complexity of the emerging systems, as in Experiment 2, did basic morphology emerge. When strict word order is not efficient enough to transmit 'who did what to whom' (agent and patient roles), speakers may adopt morphological markings to convey these semantic distinctions.

Grammaticalization theory describes how open-class lexical items may evolve into close-class grammatical items (Heine et al., 1991; Heine & Kuteva, 2002; 2007). This may also lead to 'morphologization': an independent marker becomes an affix rather than a free word. What accounts for the emergence of morphology? Our data suggest one possible reason is the facilitation of communication. In our experiment, throughout the diffusion chain, the vowels gained the specific morphosyntactic function of marking grammatical relations. Participants restructured the language and used word structure to facilitate communication. This decrease of arbitrariness supports the recent change in the notion that words have mainly arbitrary relation to its meanings. As suggested by Dingemanse et al. (2015a) a fully arbitrary language would be difficult to learn. Moreover, our results suggest that basic grammaticalization processes in vertical and horizontal transmission could be universal and not tied to a speaker's native language. Our results support the claim that while languages vary, systems of language use may be similar cross-linguistically (Dingemanse et al., 2015b).

Our results suggest that regular linguistic structures that parallel typological linguistic patterns could emerge due to communicative pressure and vertical transmission. Reali & Griffiths (2009) have shown in their computational work that a single generation of learners may not necessarily reveal a bias towards regularisation. The bias becomes evident only after several generations with cultural transmission. As in our experiment, grammaticalization was not evident in a single generation. Only after several generations did languages became regular and did proto-morphology emerged. It is important to note that our codes are a simplified and limited model of natural languages, which may be a confound. Nevertheless, the results reproduce some of the core features of a linguistic design.

3.4 General discussion

The present study examined whether and how structure in simple artificial languages emerges and stabilizes in diffusion chains where players interact in iterated signaling games. In two experiments we show that by introducing communicative pressure and feedback the emerging signaling systems are highly structured and expressive right from the first generation. We replicate earlier results (Moreno & Baggio 2015) that language transmission within each generation is largely unidirectional from senders to receivers. We show for the first time an increase in *coordination* and *transmission* across generations: along a chain it becomes easier for players in a game to agree on a common code, and the codes agreed upon by subsequent player pairs or generations become increasingly similar. In the second experiment, on the emergence of grammatical functions, we observed a decrease of *innovation* across generations, which was not seen in the first experiment.

Natural languages share certain basic features, some of which may be shaped by coordination and communication pressures (Beckner et al., 2009, Fedzechkina et al., 2012, Monaghan et al., 2014). In each game, players interacted repeatedly, with the explicit goal of arriving at a shared signaling system. As in actual communication, if interlocutors cannot reach a mutual understanding, they will keep exchanging signals until a consensus is achieved. Signaling games are effective formal, computational and laboratory models of coordination-for-communication and the transmission of codes within a pair, i.e., from sender to receiver (Moreno & Baggio, 2014; Lumaca & Baggio, submitted). We replicated previous findings (Moreno & Baggio, 2014) that, when each player is either sender or receiver throughout a game, codes tend to be faithfully transmitted from senders to receivers. Selten & Warglien (2007) suggested that role asymmetry is one of the keys of communicative success. We show this spontaneous asymmetry also applies to diffusion chains, suggesting – as a proof of concept – that signaling games are a working model to study the transmission of languages across 'generations' (i.e., multiple links in a chain). Besides being a viable laboratory model of cultural transmission, signaling games have some desirable properties (simplicity, flexibility, formal explicitness, algorithmic implementability) that allow us to connect experimental data with game-theoretic and computational models of coordination and information flow (Skyrms, 2010).

3.4.1 Communicative pressure and language change

A considerable share of our knowledge and skills is acquired through interaction with others such as by observing and imitating the actions of tutors and peers. Languages are learned by listening to others speak *and* by interacting with them. In generative linguistics and related approaches, following Chomsky, language is seen as a faculty of the individual mind. In order to account for language acquisition and competence, a genetically and neurally encoded 'language organ' has been proposed, which consists of a universal grammar that delimits the space of possible languages and constrains what can be learned (Chomsky, 1965). It is, however, still an open issue to what extent language acquisition is constrained by language-specific or domain-general 'internal' factors (Culbertson et al., 2012) and by 'external' forces, such as cultural transmission (Dunn et al., 2011; Levinson & Gray, 2012). The capacity to acquire a language is transmitted biologically, but languages themselves are transmitted culturally. Unless our biological make-up determines precisely how languages look like, and thus drives entirely the forms of variation observed historically, it is likely that certain aspects of language structure are the result of cultural transmission and of the cognitive biases and constraints that operate at the level of social coordination and interaction. That is why a theory of language structure may, at least in principle, require models of social exchange of the kind provided by game theory and related formalisms.

Observations of how languages change may contribute to current discussions of the cognitive abilities underlying language acquisition. We are only now beginning to understand how social learning may shape languages. Garrod & Anderson (1987) have shown that, while interacting, partners align their meanings through changes at different levels of form, including syntax (Branigan et al., 2000), prosody (Giles et al., 1991) and morphology (Beckner et al., 2015). Direct forms of semantic alignment, in which players adopt the same labels of events during interaction, was also observed in modeling studies (Steels, 2003; Barr, 2004). At the same time, Cuskley et al. (2014) demonstrated using corpus data that languages change in a self-organizing way as the consequence of exogenous and endogenous pressures to minimize processing effort while maintaining a sufficiently high level of expressivity. Nowak and colleagues (Nowak and Komarova, 2001, Nowak et al., 2002; Nowak, 2006) have developed mathematical model of language and cultural evolution. They explored how grammar and lexicon co-evolved as a system to facilitate cooperative relationships among social groups. They suggest that even minor improvements in the communicative efficiency are sufficient for these improvements to be spread in social groups. Moreover, they proposed that cooperation is the key to the evolution of complexity. One outstanding question is whether language change necessarily simplifies linguistic structure. Our Experiment 1 shows that when codes are simple enough (i.e., there is compositionality and linear order), there is no decrease in innovation over generations, the codes keep changing. However, in Experiment 2, where the codes were more difficult to learn, we did find a cumulative decrease in innovation, which shows that codes changed until they were perceived by players as optimal or 'good enough' in some respect.

3.4.2 Compositionality and proto-morphology

A striking property of language that differentiates it from most other communication systems in nature is that the meaning of signals is composed by meaningful subparts. The advantages of combinatoriality in languages have been highlighted by modeling studies (Nowak et al., 2001). But where does this compositional structure come from? And what maintains it in language? One answer is that compositional structures arise from cultural transmission when languages are under pressure to be both expressive and learnable (Kirby et al., 2008; Smith et al., 2013; Kirby et al., 2015). Our results suggest that communicative pressure accelerates this process, minimizing the role of vertical transmission.

It is worth comparing our results with those of Kirby et al. (2008, 2015). They showed that structure emerges as a consequence of language transmission over generations, without intentional design on the part of individuals. The authors observed convergence to a system with compositional structure, while we found that compositional structure emerged already in the first generation as the result of dyadic interaction. Bowie (2008), using restricted language systems, has shown that compositionality significantly enhances communication in novel contexts. Players in our signaling games appear to use compositionality as an anchor and gradually agree on detailed semantic associations between signals, constituents, and their referents. Whereas Kirby et al. (2008) used an artificial procedure to produce the final system (i.e., eliminate homonymy), we obtained comparable results in signaling games with communicative pressure and feedback.

In Kirby et al. (2015), the authors show that only an interplay between expressivity and learnability pressure results in structured languages. Our data are consistent with these results, with some differences. In our study we used signals that may be closer to the stimuli learners encounter when they learn a language. The messages sent by senders are sentence-like, with distinct words as in natural languages. In Kirby's experiments, it is harder to know whether learners treat the signal as a compound word, or whether they parse the message into smaller discrete units, in a manner similar to syntactic processing. The meaning of compound words can be equivalent to the meaning of its components, but this does not need to be the case: compare 'breakfast' and 'daybreak' in English (Seidenberg & Gonnerman, 2000). In our stimuli, constituent boundaries were clearly present within signals. We believe that even though this is a small difference in the visual presentation of stimuli, it may have a significant impact on parsing and generating signals, and indirectly on the kind of code change phenomena we have observed.

In Experiment 2 we have observed a split of lexical and functional morphology in the construction of grammatical categories across a transmission chain. Speakers regularized code structure and introduced basic morphosyntax. Language can be seen as a dynamic system in which individual words adapt and contribute to decoding the context in specific ways (Silvey et al., 2014; Kirby et al., 2014). In codes at the end of

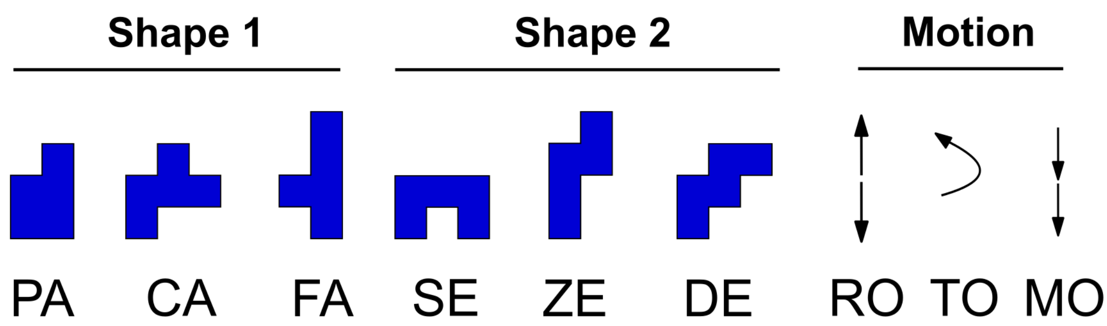


Figure 3.7: The example of final signal-referent mappings in Italian Chain 4.

a chain, vowels represented grammatical categories and consonants represented exact semantic associations within a category. Garrod et al. (2007) argued that interaction shifts the locus of linguistic information from the linguistic code itself to the learner's memory of the code's usage. Our results point to the mechanisms of how words may be remapped to meanings and grammatical categories in adaptive ways as a product of incremental changes in pair-wise interactions. The emergence of basic morphology was not found in single dyads but only after several generations.

3.4.3 Future directions

A necessary simplification in our study is that each generation consisted of only two individuals. Global coordination often takes precedence over pairwise coordination in establishing and maintaining a community wide linguistic system (Fay et al., 2010). Fay et al. (2000) have observed two different kinds of face-to-face communication in small and large groups. In small groups, they found a bilateral process of establishing consensus among pairs of communicators. In large groups, they identified a unilateral process of broadcasting information within a group. Fay et al. (2008) suggested that communities evolve more effective graphic signs than isolated pairs. However, it is not clear whether this applies to linguistic material too. We predict a lower flexibility in language change for the exact mappings between signals and referents at the group level as compared to our dyadic design. This requires further research. As one of our aims was to determine whether signaling games can perform as a model of language transmission, we focused on pairwise interactions only. We also wished to establish a baseline for future work using iterated signaling games. In further experiments, it would be useful to extend signaling games to more than two players (Skyrms 2009, 2010), and to examine how consensus is achieved and how codes change.

One obvious limitation of our study is that our participants were adults with full competence of their native language and life-long experience of social interaction. This suggests some care in interpreting experimental results. The only way to confirm and generalise our findings, as well as the results of IL research, is to conduct these studies on children. Indeed, language acquisition has to be incorporated in studies on language change to yield a more complete picture of what drives structural variation (Monaghan, 2014). There is an ongoing debate on the role of adult and child learners in language change (Croft, 2000; Hudson Kam & Newport, 2005; Slobin, 2005, Chater & Christiansen, 2010). One of the criticisms that may be addressed to the IL is that language is not acquired only from peers with comparable levels of experience (Vogt, 2005). We tried to avoid this by using 'generations' that consist of one experienced player (an 'adult') and one naive player ('offspring'). This simplification needs to be dropped in further research, where dyads comprising of an adult and a child can also be included in transmission chains. This highlights both a limitation of the present study and the flexibility and power of ISGs to progressively accommodate a variety of real world scenarios.

Chapter 4

Artificial language processing: an ERP investigation into adults' early stage learning effects of semantic and compositional constraints

Recent studies suggest that adult language learning can come to rely on native-like language brain mechanisms. How quickly are new grammatical structures integrated into the language system and under what circumstances? We approached the question by investigating the electrophysiological correlates of learning semantic and compositional constraints with a miniature artificial language. We conducted a two-day study composed of two phases: behavioral training via signaling games and EEG test phase. Brain activity was measured during the visual presentation of stimuli in either correct or incorrect contexts, which varied in two main ways: syntactic or semantic violations. The findings show that novel words and syntactic structures can be acquired quickly without explicit instructions and are integrated into language network rapidly. The ERP effects point to early stage effects of learning in both syntactic and semantic processing. The learning material influenced the final learning outcome - the structures containing syntactic manipulations yielded P600 but only if the structures were present in the learning material. Semantic manipulations elicited N400-like effects to limited type of stimuli. Overall, the results suggest that conditions under which the language is learnt may be essential at determining the neurobiology of language learning.

4.1 Introduction

Children acquire their native language at remarkable speed with minimal effort, given the complexity of natural languages. Adults can learn second or additional languages later in life and still communicate efficiently. However, it is widely believed that if a language is acquired at later stages, language processing will differ from native language processing and different mechanisms will be at work during learning (Chomsky, 1965; Weber-Fox & Neville, 1996; DeKeyser, 2000, Perani & Abutalebi, 2005; Costa & Sebastian-Galles, 2014; Abutalebi, 2015). Only if a language is acquired within a certain 'critical period' the native-like competency can be achieved (Clahsen & Felser, 2006). More recent theories posit that after sufficient exposure and learning, second language processing can operate in the native-like way (Birdsong, 1992; Christiansen & Chater, 2008). This view highlights the high levels of proficiency that can be attained by late language learners. During language learning, processing seems to undergo systematic changes and exhibits a growing proficiency level (Steinhauer et al., 2009). The claim is that language processing after acquiring it in adulthood is not qualitatively different from L1 processing (Ellis & Cadierno, 2009). This is supported by brain imaging studies showing that the same brain areas are recruited for first and second language processing (Musso et al., 2003; Indefrey, 2006) and EEG studies demonstrating that native-like brain signatures of syntactic processing can be found for artificial language learners (Friederici et al., 2002; Morgan-Short et al., 2012). However, it is still unclear whether native-like processing can always be obtained or whether certain factors such as the nature of the learning environment and learning material can constrain it. Questions concerning first and second language learning are difficult to answer in laboratory settings, because it is extremely difficult to control for the many factors that have been shown to affect language learning and processing (e.g. the amount of exposure, similarity to learners' first language, frequency). In this study we therefore turned to simple artificial languages to investigate early stages of language learning.

The Artificial Language (AL) learning paradigm is a relatively new approach for studying language learning in a controlled manner. To date most of the experimental literature has separated vocabulary and syntax learning. In several paradigms, usually lexical items are learned first (Friederici et al., 2002; Musso et al., 2003; Newman-Norlund et al., 2009) and in the next phase grammatical rules are either explicitly presented (Musso et al., 2003) or implicitly acquired (Friederici et al., 2002; Newman-

Norlund et al. 2009). Here we combine simultaneous vocabulary and rule learning through implicit learning via signalling games. Our study adds to the small number of studies that it used artificial language learning in combination with EEG.

Semantic and morphosyntactic processing in the brain are extremely fast, and often result in modulations of partly overlapping responses. Therefore, the high temporal-resolution of a technique like EEG is appropriate. Most of the studies on artificial languages or second-language learning use behavioral measures and have suggested the viability of the AL paradigm for investigating constraints on natural language learning (Culbertson, 2012). A small number of studies combining AL and EEG were able to reveal differences between behavioral performance and ERP effects. For example, Morgan-Short et al. (2012) demonstrated that behavioral performance did not differ between explicitly and implicitly trained groups on an artificial language whereas ERP measures displayed striking differences between two groups. McLaughlin et al. (2004) found that the earliest evidence of learning, even before changes in behavioral performance, leaves traces on ERP responses to novel words. This suggests that early stages of language learning may have been overlooked by current behavioral paradigms (Winkler et al., 1999; Osterhout et al., 2006; Kaan, 2007). Therefore EEG might represent a uniquely sensitive technique for exploring the developing linguistic competence of adults, and might more accurately reflect continuous change in their knowledge of the language than other measures.

4.1.1 Language-related ERPs in native speakers

Below we briefly discuss language-related Event Related Potentials (ERP) found in native speakers (for a review see Kaan, 2007). ERPs respond differently to syntactic and semantic aspects of sentences processing. The N400 was the first brain response to linguistic stimuli to be discovered (Kutas & Hillyard 1980, Kutas & Hillyard 1983). The N400 varies in amplitude as an inverse function of the degree of semantic fit between the eliciting word and the context in which it occurs (e.g. 'He liked lemon and sugar in his *soup'): the context can be either a single word (Holcomb, 1993), pictures (Federmeier & Kutas, 2001), gestures (Amoruso et al., 2013) or sentential context (Van Petten et al., 1999). Federmeier & Kutas (1999) have shown that the violations within a semantic category elicit significantly smaller effects than between-category ones, which indicates that sentential context and the long-term organisation of semantic memory interact and determine processing costs. Aspects of access into long-term memory have also been shown to alter the N400 (Kutas & Federmeier, 2000).

Left Anterior Negativities (LAN) are more frontal and left-lateralized than N400 effects (which typically show a centro-parietal distribution) with a somewhat earlier latency (in some cases it is reported to be as early as 125-180 ms) associated to violations of word-category constraints (Hahne & Friederici, 1999; Neville et al., 1991) as well as to processing of closed class words as opposed to open class words (Brown et al., 1999). If the context requires a word of a certain syntactic class and the word from a different class is presented (e.g. 'She hoped to *train catch early'), a LAN is often elicited (Hagoort & Wassenaar, 2003). Early Left Anterior Negativities (ELAN) appear to be the earliest known component of syntactic processing, with a latency around 100 to 300 ms. Its topographical distribution varies somewhat across studies, but is usually found to be maximal over frontal electrodes (King & Kutas, 1995). The ELAN has been observed while processing phrase structure violations and thus is suggested to be a purely syntactic component (Neville et al., 1991; Friederici et al., 1993; Steinhauer & Drury, 2012). It has been proposed to reflect automatic early processing during which word category information is used to build up an initial syntactic representation of phrases (Friederici et al., 1995; Friederici, 2002). However Steinhauer & Drury (2012) have claimed that so far there has not been enough evidence for a first phase being exclusively restricted to phrase structure processing.

An additional syntax-related ERP component is a late positivity with a centro-parietal distribution and positive polarity starting at about 500ms and extending up to even 800ms. Even though the presence of a clear syntactic violation is not necessary for a P600 to occur (Kaan et al., 2000), the common view is that P600 is associated with syntactic processing. P600 occurs when a word or phrase is difficult to integrate into the structure of the preceding context . P600 has been found to be elicited by a variety of syntactic manipulations and structure violations (e.g. 'The dog will *eating') (Neville et al., 1991; Hagoort et al., 1993; Osterhout & Holcomb, 1993; Rosler et al., 1993; Friederici et al., 1996; Gouvea et al., 2010) and is susceptible to differential task effects (Schacht et al., 2014). The P600 is thought to reflect syntactic integration (Hagoort, 2003) or syntactic re-analysis and repair following the detection of an ungrammaticality (Friederici, 2002; Molinaro et al., 2011). It has also been suggested that the P600 reflects domain-general (not only syntactic) re-analysis or monitoring processes (Munte et al., 1998; Kolk et al., 2003) and possibly processes of other aspects of language, such as pragmatics (Regel et al., 2014).

4.1.2 Non-native language learning and brain imaging

An open question is to what extent and when do learners employ the same or similar neural and cognitive mechanisms as native speakers. The use of brain imaging and electrophysiological techniques has shed new light on the neural bases of second and artificial language processing. The current view, supported by the ERP data, is that artificial and second language learners use qualitatively similar lexical processing mechanisms as do native speakers (Steinhauer et al., 2009).

ERP studies of non-native language learning have revealed that the neural correlates of lexical processing do not qualitatively differ from native processing, reliably yielding N400 effects. Some studies have shown that N400 effects elicited by semantic anomalies are slightly delayed or longer lasting (Ardal et al., 1990; Kutas & Kluender, 1991; Weber-Fox & Neville, 1996; McLaughlin et al., 2004; Mestres-Misse et al., 2007, Steinhauer et al., 2009) suggesting that semantic processes are somewhat slowed down. Lexical effects have been shown to emerge relatively early in second language learning (McLaughlin et al., 2004). Borovsky et al. (2010) examined the impact of the initial learning context on understanding novel words using ERPs. Their results demonstrate that rapid word learning is modulated by contextual constraints and that adults are more likely to be learning novel words via incidental or implicit learning. Even a single exposure to a new word in a highly constraining sentential context can be enough to evoke neural signatures of semantic processing (Borovsky et al., 2012). Yet other studies have demonstrated that the lexical integration may not be complete after initial exposure.

It is well established that memory performance and generalization are enhanced with sleep (Walker & Stickgold, 2006). Novel words require a post-learning consolidation phase during which neocortical connections are strengthened. Recently, Bakker et al. (2015) have shown that the difference in N400 amplitude between novel and existing words decreased significantly after merely 24 hours of consolidation period, providing support for the hypothesis that offline consolidation aids lexicalization. They proposed that the underlying semantic processes may become more automatic and native-like with consolidation.

On the other hand, syntactic processing seems to differ depending on the level of proficiency. At a lower proficiency levels, anterior negativities are generally nonexistent (Hahne & Friederici, 2001; Ojima et al., 2005; Mueller et al., 2005) or are replaced by N400-like posterior negativities (Weber-Fox & Neville, 1996; Osterhout et al., 2008). At higher level of proficiency anterior negativities have been reported by Steinhauer et al. (2009). Interestingly, also proficient bilinguals do not always display an early left anterior negativity (Weber-Fox & Neville, 1996). P600 effects have been found at higher proficiency levels (Osterhout et al., 2008) as well as native like biphasic responses (Friederici et al., 2002; Morgan-Short et al., 2012). Friederici et al. (2002) have pointed out that speaker's proficiency can tune processing to be more native-like. They obtained results suggesting that adults who learned a miniature artificial language display a real time pattern of brain activation when processing this language, similar to native speakers. Their results demonstrate that a small system of grammatical rules learned implicitly can strongly resemble native-speakers' sentence processing.

Morgan-Short et al. (2012) have established that also the type of training that learners receive impacts the outcome of the learning. They used artificial language to examine, longitudinally, whether explicit training (that simulates traditional grammar-focused classroom learning) and implicit training (immersion-like) affect neural and behavioral measures of syntactic processing in the same way. They found that implicit training yielded N400 at low proficiency, while at high proficiency a biphasic response with an anterior negativity followed by a P600 was found. On the other hand, explicit training elicited no effects at low proficiency and only an anterior positivity followed by a P600 at high proficiency. Interestingly, these differences could not be explained by behavioral measures, performance of explicitly and implicitly trained groups did not differ at either low or high proficiency.

In an fMRI study of Musso et al. (2003) participants were explicitly taught a subset of the rules of a new (to them) natural language. The participants received lists with words and learnt the real and unreal grammatical rules between fMRI sessions. They found LIFG activation to be related to the outcomes of explicitly learnt new grammatical rules. Similar result was obtained by Tettamanti et al. (2002). The authors compared the neuroanatomical correlates of learning grammatical and non-grammatical rules. Only the grammatical rules activated Broca's area. Other artificial language studies reported neural activations that were similar to those found for natural languages (Opitz & Friederici, 2003; Opitz & Friederici, 2004). They also show the involvement of the LIFG and the correlation of the activity with the proficiency level.

In summary, the studies combining AL and brain imaging techniques show that although lexical processing is similar in L1 and newly acquired languages, neural responses to syntax processing may depend on the learner's level of proficiency and type of exposure.

4.2 Aims

Our goals were twofold. Our first aim was to investigate brain responses in the initial stage of learning an artificial language and to study the neuropsychological foundations of basic compositional processing of artificial language-like structures. To this end we designed an AL learning paradigm in the form of signaling games (i.e., a two-player coordination and communication game; details below) that simulates natural language acquisition through communication. The word order rule was not taught explicitly but was to be inferred based on the regularities of the structures and accompanying feedback. We also reduced the complexity of the stimuli to a simple noun phrase (noun and adjective) and verb. The important part of the design is that subjects are learning the language always through 3-feature sentences with fixed order. Using the compositional constraint of the strings of words and their visual representation, a meaning is mapped to a novel pseudoword based on the word order.

Our second goal was to provide further validity for AL paradigms, especially learning through signaling games, and investigate whether AL processing after such a short training resembles natural language processing. There have been numerous behavioral studies exploiting the AL paradigm but only few involved EEG data. Behavioral measures are not necessarily sensitive enough in early stages of learning, and there may be neural precursors of behavioral effects. Using EEG we can assess neural activity in minimal compositional contexts, using both correct strings (i.e., conforming to the learned patterns) and strings containing violations (see below).

We used a two-day paradigm: on the first day participants learn the artificial language and are tested on it only the day after, allowing approximately 24 hours for post-learning consolidation. Recent EEG studies investigating second or artificial learning focused on either syntactic or semantic violations only. We measured activity during the visual presentation of linguistic stimuli presented in either the correct or incorrect context, which varied in two main ways: violations of word order or violation of semantic content. We wanted to explore to what extent a minimal system of grammatical rules can be syntactically instantiated by learners in a way that resembles native-speaker sentence processing.

4.3 The present study

For many years the focus of EEG studies has been on complex syntactic structures and implausible sentence completions in the context of full sentences or discourse. Paradigms that used linguistic material with different types of violations were preferred rather than those that looked at elementary specific operations (e.g. word order, adjacent dependencies, movement). Direct investigations into the neural underpinnings of basic compositional constraints have been very few. One study that explored the minimal composition was due to Bemis & Pylkkanen (2011). They investigated, using MEG, neural circuits underlying a minimal linguistic composition of a simple adjective-noun phrase. Their manipulations did not reveal any increases in activity within more traditional areas during basic combinatorial processing. Instead they found increased activity in the left anterior temporal lobe (LATL) and ventromedial prefrontal cortex (vmPFC). Previously, these regions have been shown to be linked with syntactic and combinatorial processing in more complex contexts. They suggest that future work should aim at firstly building the foundation of simple linguistic operations (such as compositionality) to complete the thorough understanding of more complex linguistic operations (e.g. embedded hierarchical structures) by reducing the scope of investigations to a more fundamental level.

In our study we investigate the acquisition of basic compositional structures in the context of artificial language. We designed a miniature language composed of 9 words and one fixed word order with the underlying structure Noun-Adjective-Verb. Meaning was strictly dependent on the meaning of the words and on the structure of the strings. The learning outcome and the syntactic and lexical processes were measured approximately 24h after a training session. It is essential to mention that our structures are a very simplified and limited model of natural language. Since we are interested in the first stages of language learning, our structures bore very general properties of language, such as compositional structure. The miniature language designed for this study is a simplified one, however it shares some key features and structural organization with natural languages.

The majority of EEG studies on language have focused on the characterization and comparison of brain patterns resulting from many years of exposure, immersion and experience in the native language of the speakers. Equally important is the characterization of the early stages of learning, both in the first and second language, and investigating brain patterns present in late language learning. It has been shown that the learning paradigm can influence this learning outcome (Norris & Ortega, 2000). In most of the studies that combined language learning with neuroimaging the participants were exposed to the grammatical structures of the new language only after the words were acquired (Friederici et al., 2002; Musso et al., 2003; Newman-Norlund et al., 2006). This learning resembles a less natural learning style, whereby different parts of the language are acquired separately. Importantly, the majority of the natural vocabulary in adulthood is not acquired explicitly but are extracted and mapped to the forms from the context. Implicit learning may be better suited for mimicking the first language acquisition, as it has been shown that immersion-like learning of second language is superior to classroom teaching (Mueller et al., 2005). Fletcher et al. (2005) have demonstrated in the fMRI study that explicit attempts to learn the sequences produced a failure of implicit learning that represented a suppression of learning itself.

To create a more natural learning environment, characterized by the need to communicate rather than to explicitly learn rules, we employed signalling games (Skyrms, 2010). Participants learned the new language through implicit training, in which they were exposed to 3-word sentence-like strings and had to extract not only the word order rule but also the meaning of the novel words. We therefore tried to simulate natural language learning by presenting words in context, and never in isolation: the meaning of words had to be inferred based on feedback provided to participants.

Up until recently, the majority of EEG studies reporting syntactic or semantic effects, the violations are introduced in the sentence-final position. These positions in sentences are often strong attractors of global processing and are more likely to elicit N400 regardless of whether the violation is syntactic or semantic in nature (Hagoort et al., 1993; Osterhout & Holcomb, 1993). Osterhout (1997) demonstrated that anterior negativities are more likely to be present when the violation occurs at the final position of the sentence compared with the embedded violation. Thus we measured brain activity after introducing violations at every position of the string to avoid the sentence-final effect. The design of our study gave us a chance to look at the acquisition of novel words within the structural context and how meaning gets associated with the pseudowords.

Furthermore, Van Petten & Kutas (1991) have reviewed the data on N400 and concluded that it is sometimes inversely related to the frequency with which words appear in natural language. This relationship was even found to hold in isolated words. Due to using AL we also controlled for the frequency of the items (separately for the training and test phase). Moreover, other factors known to affect N400 such as length, concreteness and amount of lexical neighbours (Holcomb et al., 2002) were controlled for in our experiment.

4.4 Methods

4.4.1 Participants

31 participants were recruited for the training phase on Day 1. 5 of them we not able to finish the training phase and were thus not admitted to the EEG phase. All of the participants who finished the training on Day 1, finished the training on Day 2. 26 right-handed Italian native speakers participated in the EEG phase on Day 2. 6 participants were discarded either after the EEG session due to equipment malfunction or after preliminary data analysis due to an excessive number of artifacts in the EEG signal. 20 participants (mean age 23.7, age range 21-30, 9 female) were included in the final data analysis. They all had normal or corrected-to-normal vision, no known history of neurological or language impairments and all were trichromats. They were recruited via an institutional website and were monetarily compensated for their participation. At the end of the experiment, participants were debriefed about the goal of the study. The study was approved by the Ethics Committee of SISSA.

4.4.2 Stimuli

We used the same 3-by-3 stimulus design as in Kirby et al. (2008). The states were visual scenes varying in 3 dimensions, shape, color and motion (Fig. 4.1) as well as varying in the number of dimensions (from 1 to 3). Kroll and Stewart (1994) suggested that in early language learning stages the link between a new word and the corresponding word in the first language is more robust than the link between the new word and the concept it denotes. Hence we used artificial tetris-like shapes and ambiguous colors to prevent participants from transferring partial mappings of labels to meanings (e.g. 're' to red) from their native languages, which would act as undesirable 'focal points'. All object shapes had the same number of constituent squares (5). All motion trajectories (straight, curved and zig-zag) of objects started and ended at the same two locations on the screen. The constituents of signals were 9 monosyllables. We used closed sets for signals and states.

4.4.3 Types of trials

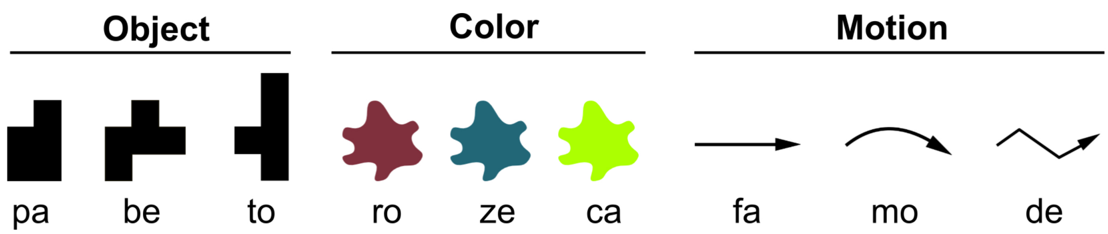
In the EEG experiment we aimed to isolate neural activity associated with basic composition mechanisms at work during the comprehension of simple linguistic phrases. We measured activity during the visual presentation of linguistic stimuli presented in either congruent or incongruent conditions, which varied in two ways: violations of phrase structure (word order) and violations of semantic content. Word order violations were created starting from correct sentences and changing the order of elements, without replacing lexical items. Semantic violations were created from each of the correct strings by replacing the last word in the string from one of the three categories with a word from either the same category (within-category replacement) or a different category (between-category replacement). Our task controlled for the length of the structure, we introduced three types of trials: one-word, two-word and three-word strings (Fig. 4.1). Basic composition was present in two-word and three-word conditions. There was no composition involved in one word trials, so only semantic replacements were used. One-word trials allowed us to assess (1) the learning effect of the new state-signal associations and (2) whether changes in activity observed during two- and three-word trials could be partially due to the presentation of multi-word strings. Compositional three-word strings consisted of a color word, a shape word and a motion word, while two-word strings consisted of two out of the three possible features (e.g., only shape and color).

In one-word trials two anomalous conditions were tested: within-category violations (e.g. instead of a word denoting Object 1, participants saw the word for Object 2) and between-category violations (e.g. instead of a word denoting Object 1, the word denoting Color 1 was presented). In two-word trials, on top of two types of lexical violations we added a third condition– scramble – in which we introduced a violation of phrase structure; lexically correct words were presented in the incorrect order relative to the one learnt in the training phase (e.g. Object-Color or Motion-Object). Importantly, in the training phase, participants were never exposed to two-word phrases. Instead, the learning material consisted entirely of three-word phrases. Finally, in the three-word trials, we introduced three main types of violations: lexical replacements (lexical substitution of the correct word), displacements (strings with correct lexical items that preserve one of the adjacencies and violate the second one) and scrambles (strings with correct lexical items, but violations of two adjacent dependencies) (Fig. 4.1). This resulted in 8 types of trials for the three-word conditions: correct (Object-Color-Movement), within-category replacement, between-category replacement, displacement where the adjacency between noun phrase and verb was violated (Motion-Object-Color), displacement with the violation of the noun phrase (Color-Motion-Object) and two types of scramble with the violations of two adjacent dependencies (Color –Object –Motion and Object-Motion-Color).

Our aim was to assess neural activity within minimal composition contexts, with the expectation that two- and three- word trials should elicit activity related to syntactic processing of linguistic composition and to investigate which types of violations would cause the most processing difficulty.

4.4.4 Procedure

To make sure that participants learnt the artificial language well enough, the training phase took place on two consecutive days and only subjects who learned the miniature language well enough (see below) were admitted to the EEG phase.



One word	
Correct	Pa.
Within-category replacement	Be.
Between-category replacement	Mo.

Two words	
Correct	Pa ro.
Within-category replacement	Pa ze.
Between-category replacement	Pa mo.
Scramble	Ro pa.

Three words	
Correct (Object-Color-Motion)	Pa ro fa.
Within-category replacement	Pa ro mo.
Between-category replacement	Pa ro ze.
Displacement 1 (Motion-Object-Color)	Fa pa ro.
Displacement 2 (Color-Motion-Object)	Ro fa pa.
Scramble 1 (Color-Object-Motion)	Ro pa fa.
Scramble 2 (Object-Motion-Color)	Pa fa ro.

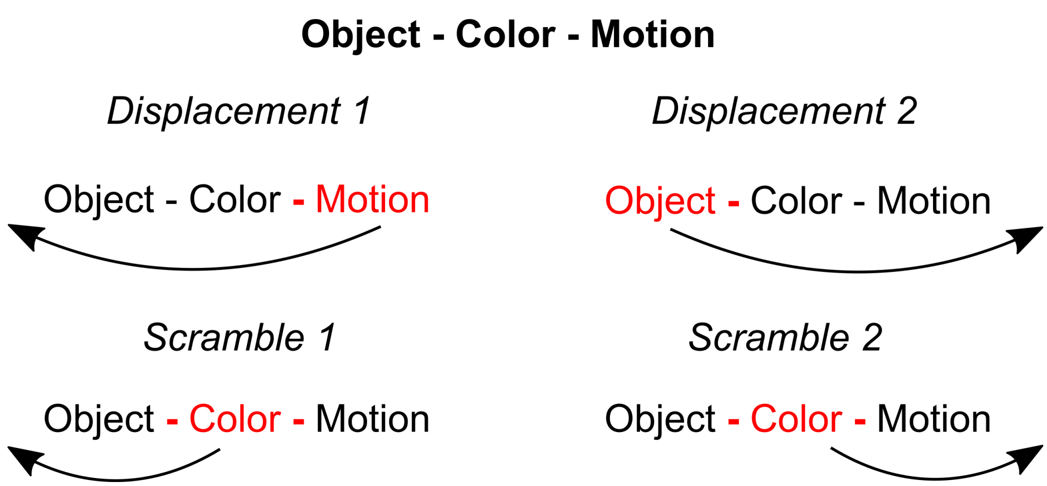


Figure 4.1: Example conditions: (upper graphs) example of associations between states and signals; (middle and lower graphs) example of semantic and syntactic violations used in the experiment

4.4.4.1 Training phase

The experiment took place on two consecutive days. On Day 1 participants' color vision was evaluated to exclude possible color vision impairment. Next, the participants performed a Training Phase. Each participant was trained using a randomly generated bijective mapping of pseudowords to object features.

The training had the structure of a Signaling Game with fixed roles: the computer acts as the sender while participant plays a receiver whose task it is to decode the signals he receives from the computer. Each trial unfolds as follows (Fig. 4.2a): the participant is presented with a three-word string (1.5s) drawn randomly from 27 possible combinations of words denoting color, shape and motion and is asked to respond by composing the event he believes the string describes. To do so, he chooses color, shape and motion features from a 3-by-3 grid showing all the options at the same time on the screen (duration: self-paced), in a random order in each trial. Following this, feedback (2s) is presented indicating whether the elements that the participant has chosen match the elements of the word string. The training ends when participant reaches 60 correct trials and has at least 2 correct responses for each event, with no constraints on the number of correct consecutive trials. A trial is correct if and only if all three features selected by the participant match the features of the event. The events presented are all equiprobable.

Importantly, the word order that all the participants were exposed to was Object-Color-Motion. To mimic implicit learning, participants were not informed about the exact word order and thus, based on the feedback, had to deduce the order of the elements. We did not have time limit for the training phase, however if after approximately an hour of training, one was not reaching 10 correct trials in total, they were not admitted to the EEG phase.

Only participants that had finished the training phase were admitted to the EEG phase of the experiment. The training phase lasted approximately 1 hour. On Day 2, participants completed a short training in order to ensure that they learnt and remembered all the associations from the day before. The procedure was the same as in the training on Day 1. However this time, in order to complete the training, participants had to provide 20 correct responses in a row. On average this took around 10-15 minutes.

4.4.4.2 EEG phase

Participants were tested in a dimly-lit sound-attenuating booth. They were seated in a comfortable chair and were informed that they were going to see a clip and then a

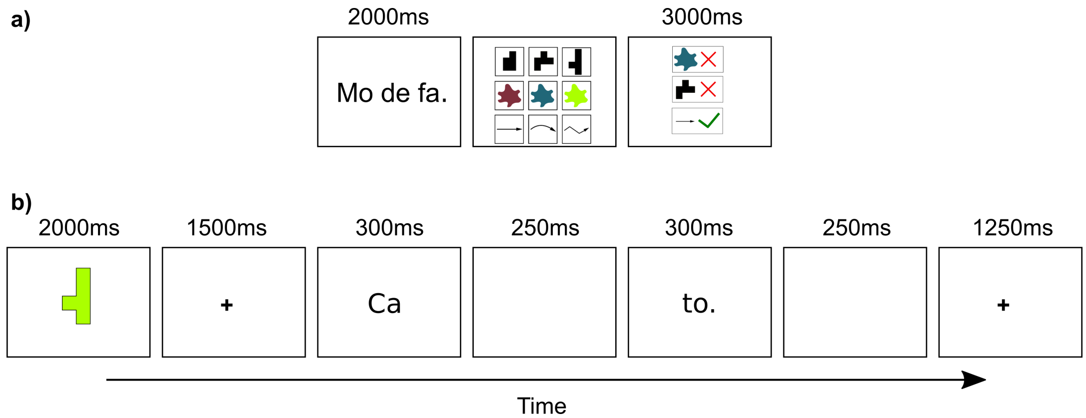


Figure 4.2: (a) An example of a trial in Training Phase. In each trial, the participants sees the signal and chooses from a 9-object array the elements that constitute the event he believes is denoted by the signal. Next, feedback is presented. Over trials, participant learns associations between signals and states. (b) An example of a trial in the EEG Phase. In each trial, the participant sees a clip, followed by a fixation cross. Next, he sees the first word of the string marked with a capital letter, followed by subsequent the words in the string.. The final word of the string ends with a period. After the final word an asterisk appears for 1250 ms. Time flows from left to right.

'sentence' that would be presented word-by-word in the middle of the computer screen. They were instructed to pay attention to the scenes and to read the sentences carefully. They were asked to try not to move or blink during the presentation of the sentence. The task therefore was passive visual exposure to visual scenes and linguistic sequences denoting them.

Each trial unfolds as follows (Fig. 4.2b). The participant sees a short animation (1.5s clip) drawn randomly from 14 types of conditions (63 clips in total), followed by a fixation cross. Next, the first word of the string started with a capital letter, the rest of the words were presented in white lower case against a dark grey background in the center of the computer screen. Each word was presented for 300 ms followed by a blank screen for 50ms. Words were separated with a 50 ms interval of blank screen to establish identical baselines and to avoid the potential silent articulation between words while still allowing for a relatively natural pace of reading. The final word of the sentence was marked with a full stop. After the final word, an asterisk appeared for 1250 ms, indicating to the participants that they could blink and move their eyes. There was a 200 ms blank interval between the asterisk and the start of the next trial.

1161 trials were presented in 5 blocks, separated by self-paced rest periods. The correct conditions contained more trials to avoid over-exposure to too many anomalous trials, thus we had double number of correct trials compared to trials with manipulations. Overall, there were 108 trials in 1-word, 405 in 2-word and 648 (e.g. 162 for correct 3-word strings and 81 for each anomalous condition) in 3-word condition-sTo ensure attention to the task, participants were presented with a question, asking whether the last string that they saw was correct or not, which appeared randomly once every 20-30 trials. Responses were made by pressing a button with the index finger of either the left or right hand. Viewing distance was approximately 90cm. The experiment lasted approximately 100 minutes and was run using Presentation Software (Neurobehavioral Systems).

4.4.5 EEG Recording

The EEG was recorded from 128 electrodes using a BioSemi system. Instead of a ground channel, BioSemi employs two electrodes, a Driven Right Leg (DRL) and a Common Mode Sense (CMS) channel, driving the average potential close to the amplifier AD-box reference voltage. DRL and CMS were placed at symmetric side positions relative to the mid-point between A1/Cz and A19/Pz. An average reference was used during the recordings. The sampling rate was 1024 Hz. The data were high-pass filtered at 0.1 Hz and low-pass filtered at 256 Hz. All filtering was digital.

4.4.6 EEG Analysis

Data were analyzed using FieldTrip (Oostenveld et al., 2011), a MATLAB package. The critical words were segmented using a window which started -200 ms before and ended 800 ms after the critical word. Segments were baseline corrected using data from the -200 to 0 ms prestimulus interval. Trials containing eye artifacts or voltage variations at any electrode exceeding 100 μV were rejected. The signal was filtered with a pass band of 1-30 Hz. Only relevant trials were selected for further analysis. The number of rejected trials in different conditions did not exceed 15% (with no asymmetry between conditions). ERPs were computed by averaging over artifact free epochs from each trial from each condition for each participant separately. Finally, grand-average ERPs were computed by further averaging over participant specific averages. Statistical analyses of ERP effects were based on a nonparametric randomization procedure (Maris

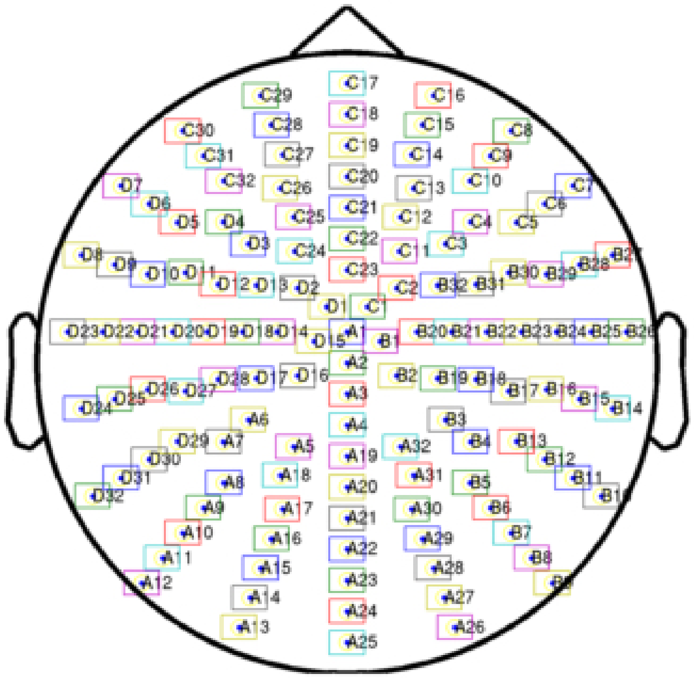


Figure 4.3: Electrode configuration used in the study.

& Oostenveld, 2007; Oostenveld et al., 2011). Firstly, participant ERP averages were compared between correct and anomalous conditional from each channel and time point with dependent samples t-tests. Next data from neighbouring time points and channels in which p-values were smaller than 0.05 were clustered together and the cluster-level t-statistics was computed from all samples included in the cluster. Finally, the cluster-level p-values were estimated using Monte Carlo simulation. This was done by creating a single set composed of participant specific ERP averages across all samples in a cluster from correct and anomalous conditions. This set was then randomly partitioned into two subsets of equal size and these subsets were compared by means of a t-test. This was repeated 1000 times. Cluster-level p-value was calculated as the proportion of partitions that yielded larger t-statistic than in the observed data.

4.5 Results

4.5.1 Training phase

The average number of trials participants needed to finish the training phase was 122.3 trials (SD = 28.8). On the second day of experiment (training before the EEG phase) participants needed on average 27.8 trials (SD=7.3) to finish the training and reach 20 correct consecutive trials.

4.5.2 Behavioral results

Participants were efficient in judging the correctness of the strings. The error rate was 19.4% (SD=7.8), this guaranteed that subjects actively attended and and processed the stimuli.

4.5.3 EEG results

For the analysis, firstly we collapsed all the lexical violation conditions together (within-category violation, between-category violation and both violations together) regardless of the number of words in the strings to see whether there was a more general effect of lexical replacement regardless of the sentential context. Secondly, we brought together all the word order anomalous conditions (4) in three-word strings and analyzed them for each word separately (first, second and third word) to investigate whether there was an effect of the word position, context or cognitive load that is more general and not tied to a particular syntactic violation. Finally, we analysed the data for all words in all the conditions (separately for each violation and different length strings), 32 in total. Below we present the statistically significant results of the nonparametric randomization procedure. Table 4.1 shows an overview of the statistics.

Firstly, we looked at the lexical manipulations that were always introduced at the last word of the strings. We found no significant differences in ERP modulations for neither within- nor between-category manipulations. As the next step we combined the data from two types of lexical replacements (between and within category) for all types of strings (1, 2- and 3-word strings). Figure 4.4 shows the grand-average waveforms of

	Time latency (ms)	t-statistics	p-value	Cluster size
Collapsed				
Lexical replacement	395-525	5.71	0.003	2689
Three words collapsed - First word	413 - 800	3.04	0.002	10048
One word				
Within-category replacement	462-663	-7.09	0.002	2881
Between-category replacement	410-584	6.89	0.01	2696
Two words				
Scramble				
Second word	220-798	-3.44	0.004	7928
Three words				
Displacement 2				
First word	300-530	8.262	0.015	2926
Scramble 1				
Second word	100-350	-1.134	0.04	3269
Third word	423-550	-2.32	0.004	1729
Scramble 2				
Second word	180-290	3.4	0.02	2464
Third word	450-560	-4.33	0.03	1892

Table 4.1: Summary of cluster-based permutation t-statistics for the ERP data comparing Incorrect and Correct conditions. The table shows the latency of clusters, T-statistics in each cluster,Monte Carlo p-values and cluster size.

the incorrect conditions and their correct controls at 9 electrode sites, as well as topographical distribution of the effect of lexical manipulation. These anomalies showed a P300-like effect: the nonparametric cluster based analysis revealed a significant positive cluster in the time window between 395 and 525 ms. P300 has been reported in the case of infrequent and task-relevant stimuli (Donchin, 1981). The topographical

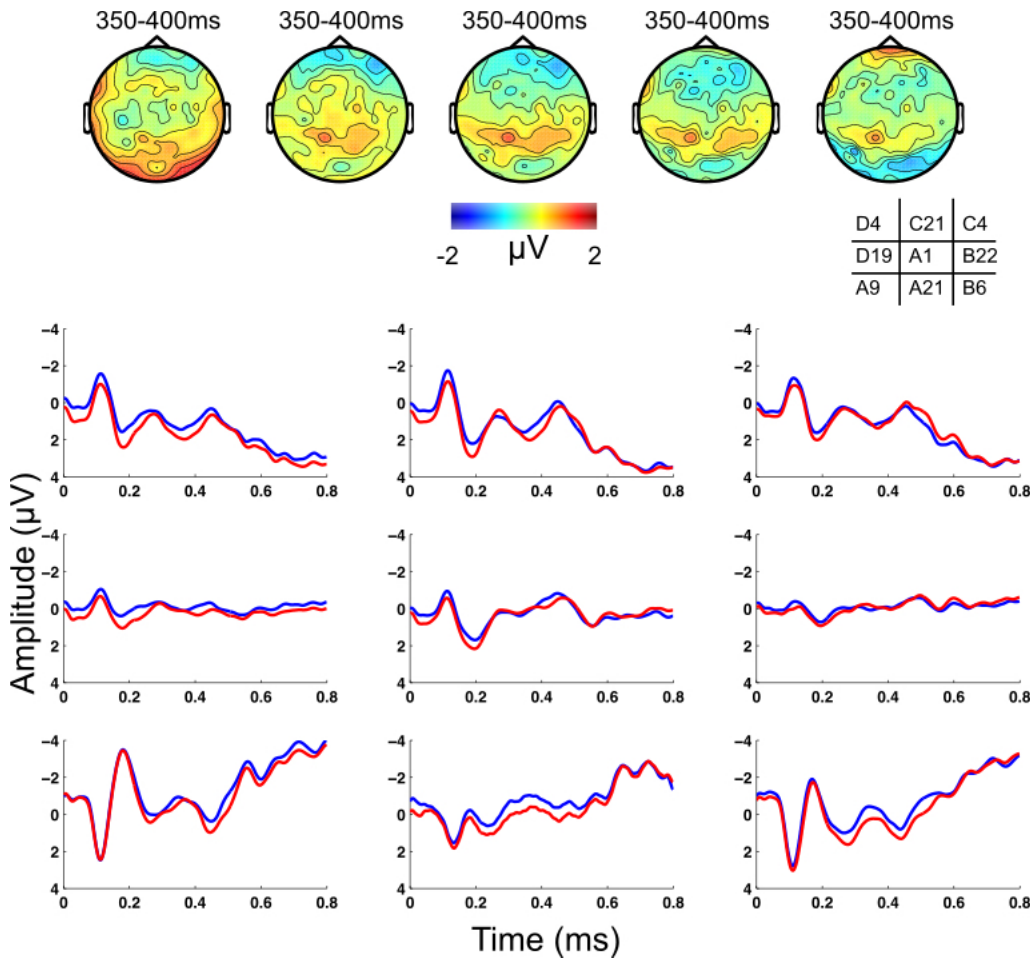


Figure 4.4: Grand-average (N=20) topographies displaying the mean amplitude difference between the ERPs evoked by two types of lexical replacement for all types of strings (1, 2 and 3-word) and their correct controls in the time window of statistically significant cluster. Lower graphs represent grand-average (N=20) waveforms from frontal, central and parietal electrode sites time locked to the onset (0ms) of the lexical violations (red) and their correct controls (blue). Negative values are plotted upward.

distribution shows the effect is strongest over posterior electrodes. This is consistent with previous studies (Osterhout et al., 1996; Evans et al., 2011).

Next we analysed lexical manipulations for different lengths of strings separately. Interestingly, we found significant differences between anomalous words and their control words only in the one-word conditions. Figure 4.5 shows the grand average waveforms of the lexical manipulation within the category in the one-word strings and their correct controls. These manipulations showed a negative effect in the time window

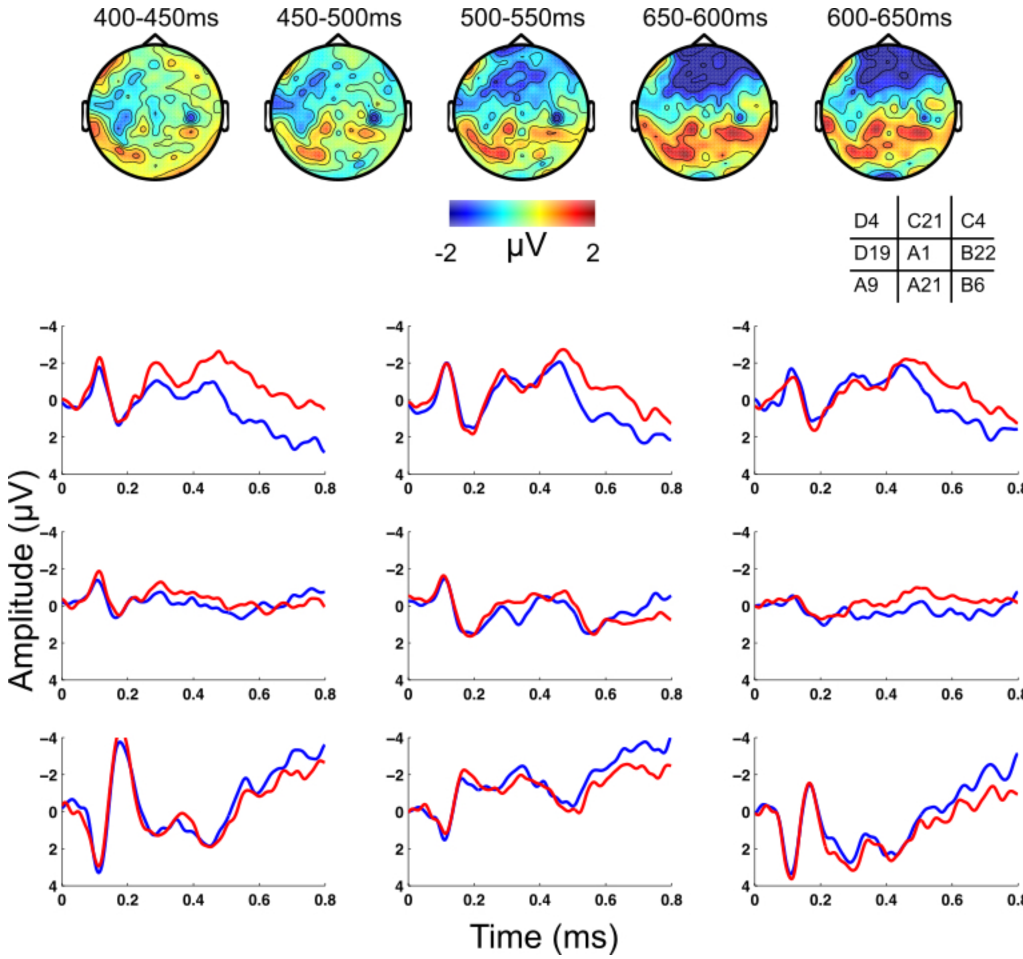


Figure 4.5: Grand-average (N=20) topographies displaying the mean amplitude difference between the ERPs evoked by within-category lexical replacement for one-word strings and their correct controls in the time window of statistically significant cluster. Lower graphs represent grand-average (N=20) waveforms from frontal, central and parietal electrode sites time locked to the onset (0ms) of the one-word within-category lexical replacement (red) and their correct controls (blue). Negative values are plotted upward.

between 462 and 663 ms. This negativity seems to be maximal at frontal electrodes and extends towards more central electrodes. A visual inspection of ERP waveforms elicited by the incorrect word reveals a negative peak around 450ms after the word onset.

Figure 4.6 presents the results of between-category lexical replacements in one-word strings revealing the posterior distribution of the P600-like effect, the topographical distribution is comparable to the standard syntactic manipulation. The cluster was present from 410 to 584 ms after the onset of the word. The waveforms reveal a positive shift starting at around 200 ms with a more pronounced positive peak in the posterior electrodes at around 500 ms. Together these results suggest that participants learned the words of the miniature artificial language. Moreover the lexical manipulations within the category (e.g. instead of word denoting Color 1, the word for Color 2 was presented) elicited a negative going effect, while semantic manipulation between-category (e.g. instead of word denoting Color 1, the word for Motion 1 was presented) elicited a positive shift. It suggests that the brain is dissociating two types of violations for words presented in isolation after having been learned in sentential context. The lexical manipulations we introduced could be treated as functional category manipulations, considering that the time window of the effect is similar for within and between manipulations, with the polarity being different.

Figure 4.7 presents the results of scramble condition in two-word strings for the second word of the string revealing the central distribution of the ERP effect. The waveforms of the incorrect words are more negative than their controls. This negative shift of the waveforms of the incorrect word was long lasting. This was confirmed by the cluster analysis: the cluster was present from 220 to 798 ms after the onset of the word. Interestingly, we found a significant cluster only at the second word, even though the violation was present already at the first word of the string (e.g. correct string: Pa ro, incorrect string Ro pa). This could indicate the effect of the training phase in which participants were never exposed to two-word strings, but instead their learning material contained only three-word strings. We did not find any other statistically significant clusters for other types of violations in two-word strings.

Finally, we analysed the data from three-word strings. Firstly, we collected all words in which violations occurred in three-word strings into three subgroups relative to the position of the word in the string, thus resulting in three conditions for the analysis: first, second and third word. We found no differences between correct and incorrect conditions in the ERP modulations for the second and third words. Figure 4.8 shows the results for the first word in three-word strings (the conditions included are Displacement 1, Displacement 2 and Scramble 2) revealing the central distribution of the P600-like effect. The topographical distribution is comparable to the standard syntactic manipulation. The positivity of the waveforms of the incorrect word was long

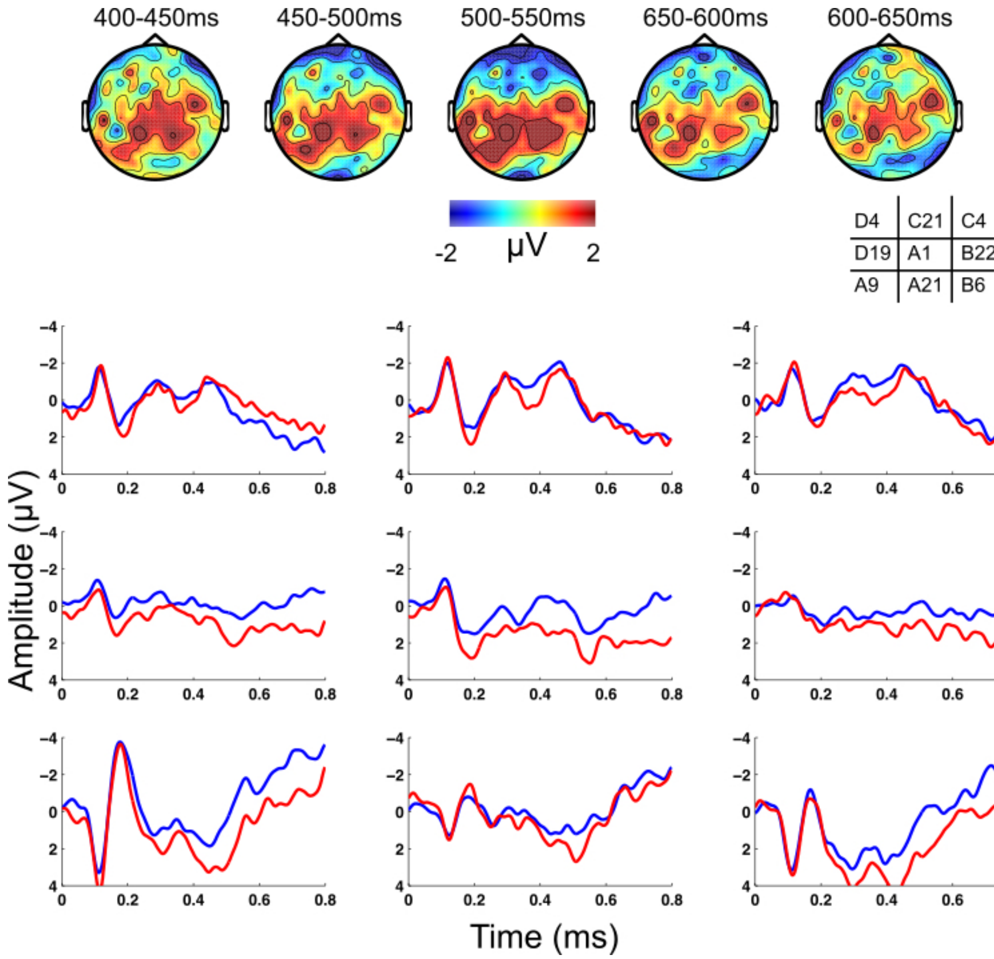


Figure 4.6: Grand-average (N=20) topographies displaying the mean amplitude difference between the ERPs evoked by between-category lexical replacement for one-word strings and their correct controls in the time window of statistically significant cluster. Lower graphs represent grand-average (N=20) waveforms from frontal, central and parietal electrode sites time locked to the onset (0ms) of the one-word between-category lexical replacement (red) and their correct controls (blue). Negative values are plotted upward.

lasting, the cluster was present from 413 to 800 ms. One possible explanation of why the only significant cluster was present for the first word could be that participants learned the structure more holistically. They did not necessarily need a context to evaluate the correctness of the string. It could also explain why we did not find significant differences for the second and third words: once the first word was evaluated as incorrect, the learners changed their processing of the string to be less holistic, which is supported by the further results.

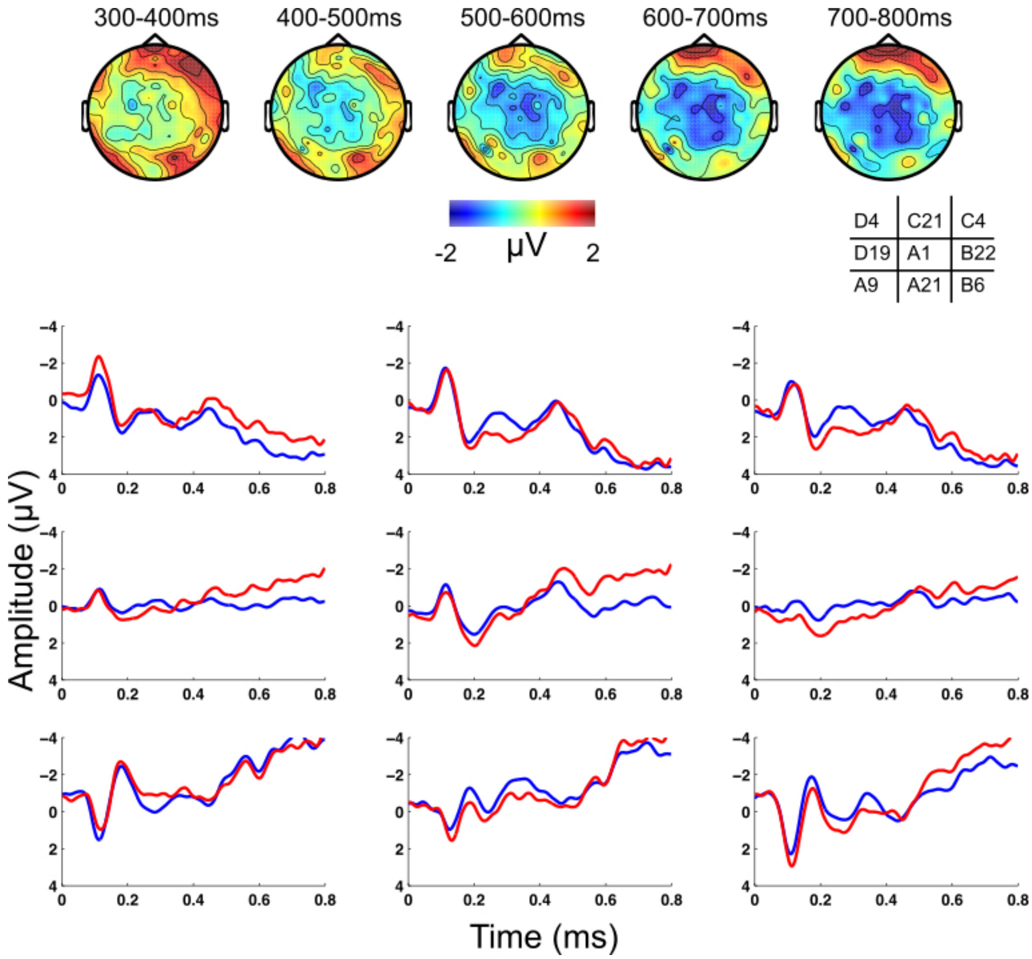


Figure 4.7: Grand-average (N=20) topographies displaying the mean amplitude difference between the ERPs evoked by the Scramble (incorrect word order) in the second word for two-word strings and in their correct controls. Lower graphs represent grand-average (N=20) waveforms from frontal, central and parietal electrode sites time locked to the onset (0ms) of the second word in two-word strings in Scramble (red) and their correct controls (blue). Negative values are plotted upward.

Our next step was the separate analysis of every word in which the violation was present for every condition separately, 32 in total. Here we present the results with significant differences between violations and their correct controls. Figure 4.9 shows the grand average waveforms of the Scramble 1 condition for the second word (Color-Object-Motion order instead of Object-Color-Motion), as well as topographical distributions of the effect. It reveals widespread distribution, being the strongest over the central electrodes. The negative cluster was present from 100 to 350 ms.

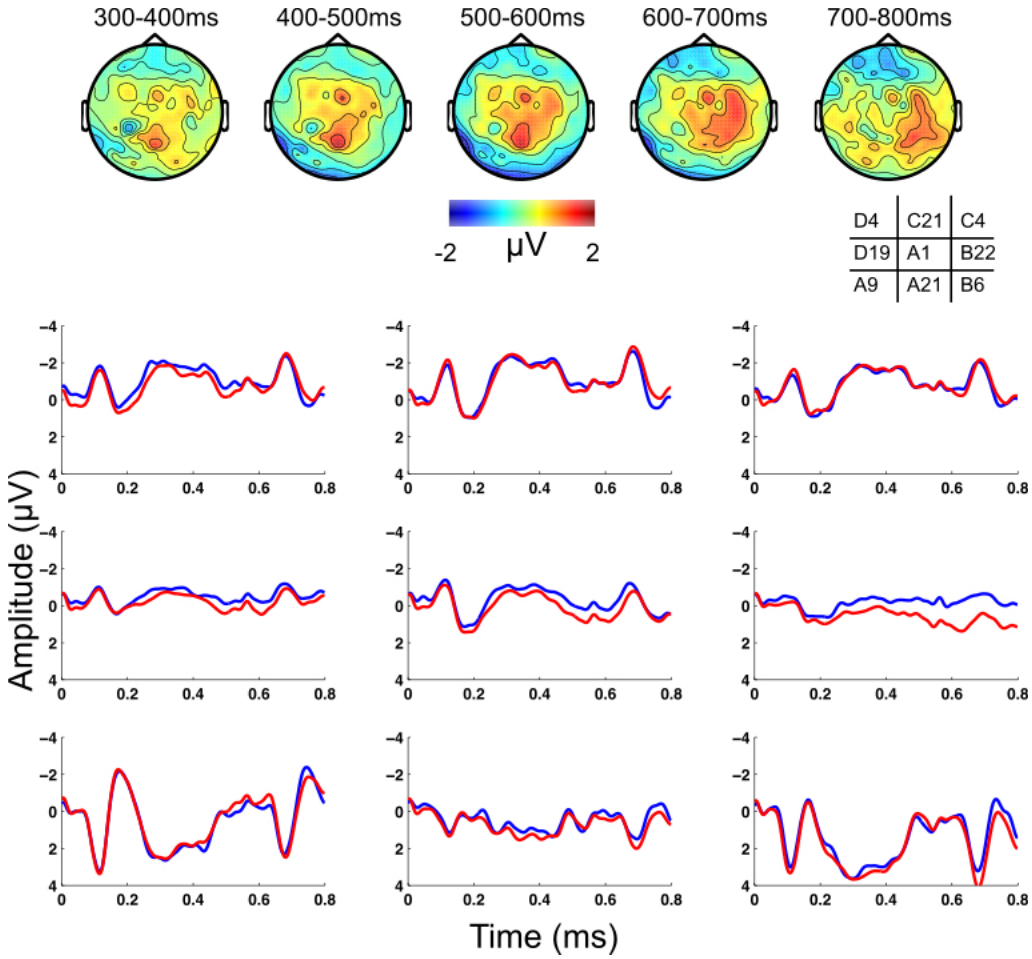


Figure 4.8: Grand-average (N=20) topographies displaying the mean amplitude difference between the ERPs evoked by the syntactic violation on the first word for all three-word strings in which first-word violation was present and in their correct controls.. Lower graphs represent grand-average (N=20) waveforms from frontal, central and parietal electrode sites time locked to the onset (0ms) of the syntactically anomalous first word in three-word strings (red) and their correct controls (blue). Negative values are plotted upward.

The second significant cluster for the second word in three-word strings was found for the Scramble 2 (Object-Motion-Color). The results are presented in the Figure 4.10. The short lasting positive cluster was present from 180-290 ms, with a posterior distribution. Interestingly, Scramble 2 is the only condition within the three-word strings where the first violation appears for the first time at the second word. This could explain the positive amplitude which could be compared to the effect present in the first word for all syntactic violations, however less pronounced and shorter.

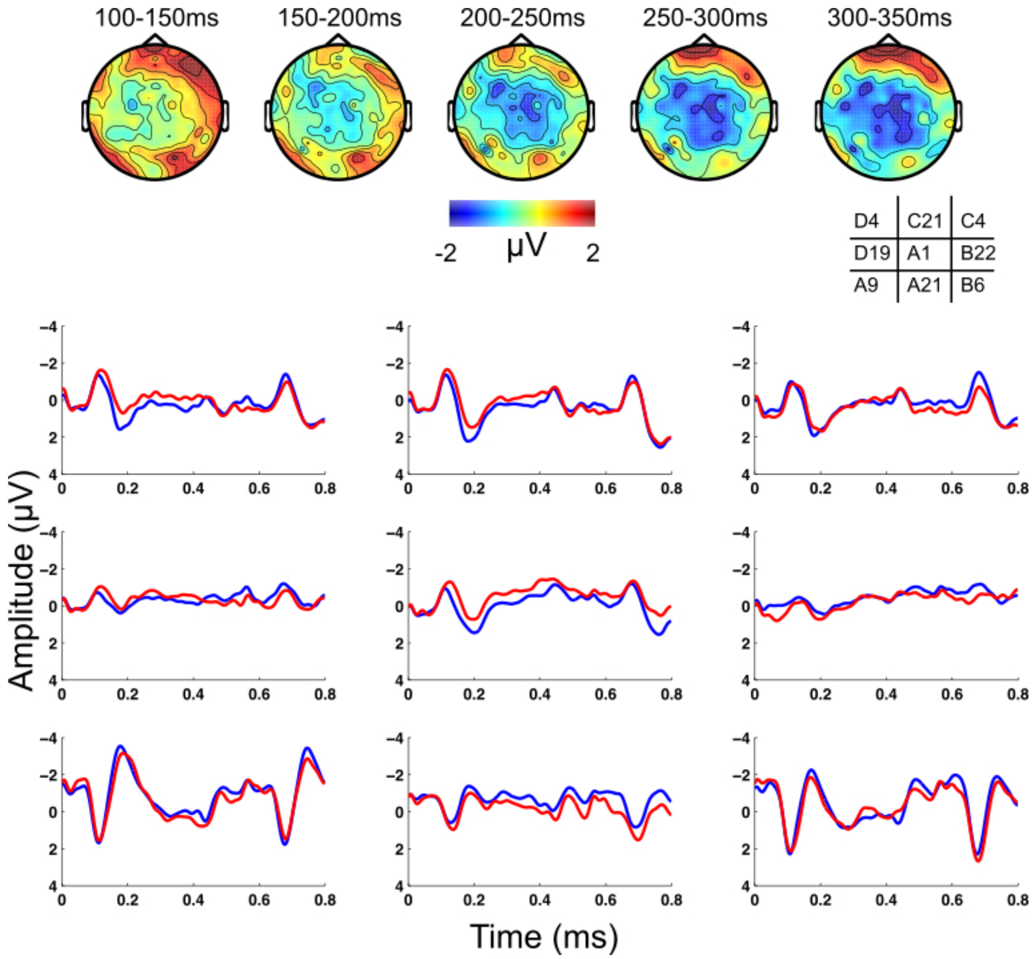


Figure 4.9: Grand-average (N=20) topographies displaying the mean amplitude difference between the ERPs evoked by the syntactic violation on the second word in three-word string in Scramble 1 and in their correct controls. Lower graphs represent grand-average (N=20) waveforms from frontal, central and parietal electrode sites time locked to the onset (0ms) second word in Scramble 1 (red) and their correct controls (blue). Negative values are plotted upward.

Finally, for the words in the final positions of the three-word strings we found significant differences in two conditions: Scramble 1 (Color-Object-Motion) and Scramble 2 (Object-Color-Motion). The time window for the two negative clusters was almost identical, 423 to 550 ms for Scramble 1 and 450-560 ms for Scramble 2, as well as a topographical distribution that shows the strongest effect over the central electrodes. Figure 4.11 shows the grand-average waveforms and topographies for Scramble 1. Visual inspection of ERP waveforms reveals a negative shift at around 220 ms after the

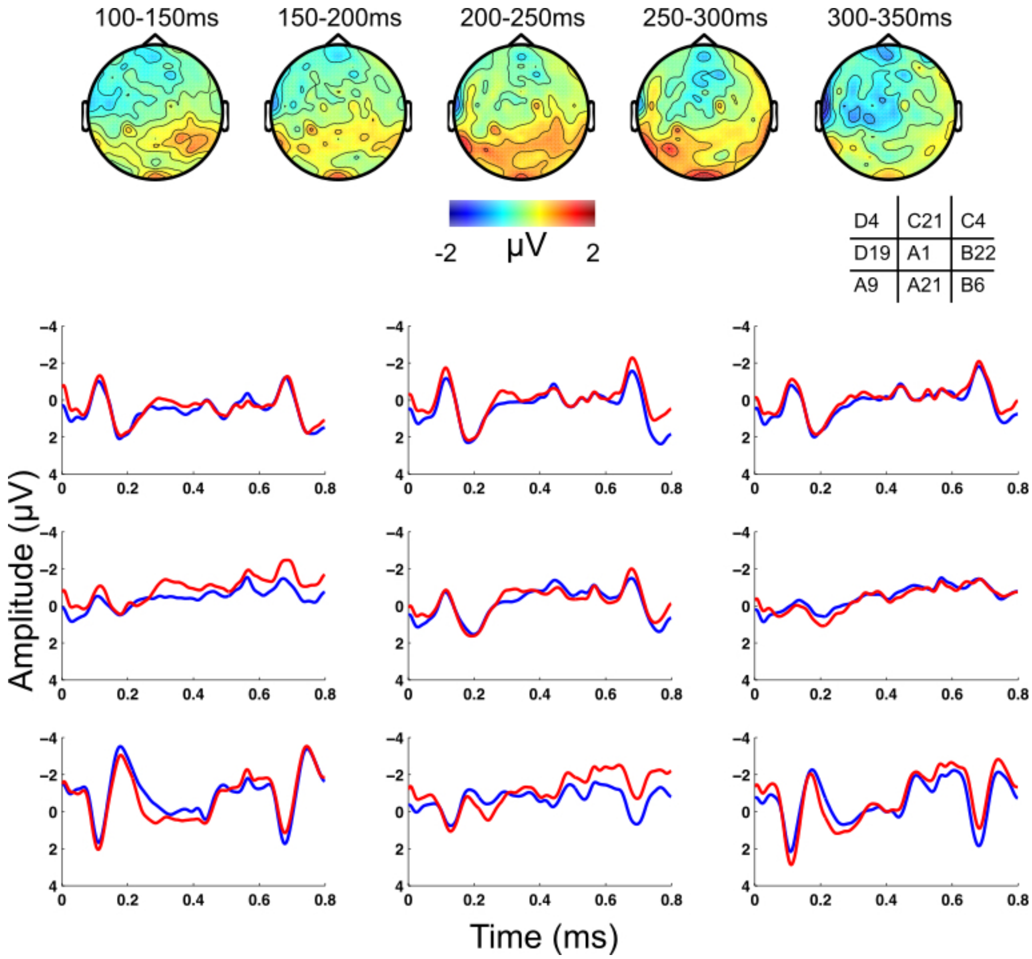


Figure 4.10: Grand-average (N=20) topographies displaying the mean amplitude difference between the ERPs evoked by the syntactic violation on the second word in three-word string in Scramble 2 and in their correct controls. Lower graphs represent grand-average (N=20) waveforms from frontal, central and parietal electrode sites time locked to the onset (0ms) of second word in Scramble 2 (red) and their correct controls (blue). Negative values are plotted upward.

word onset, peaking at around 500 ms (A1 electrode). A similar effect was found for the third word in Scramble 2 (Figure 4.12). The difference between violation and its correct control emerges around 200ms and reaches its peak at around 480 ms. The word in a Scramble 2 elicited a larger N400 compared to the word in the control condition. Figure 4.13 shows significant clusters for all the conditions.

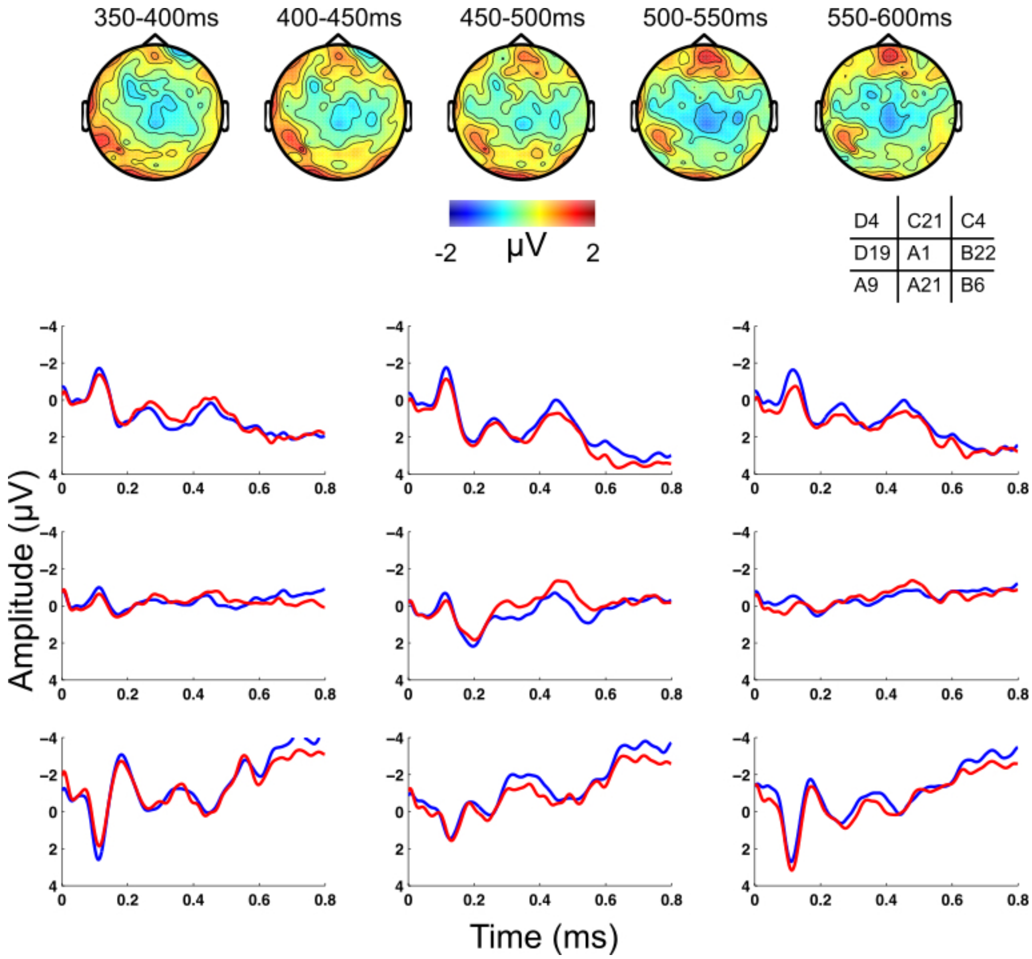


Figure 4.11: Grand-average (N=20) topographies displaying the mean amplitude difference between the ERPs evoked by the syntactic violation on the third word in three-word string in Scramble 1 and in their correct controls. Lower graphs represent grand-average (N=20) waveforms from frontal, central and parietal electrode sites time locked to the onset (0ms) third word in Scramble 1(red) and their correct controls (blue). Negative values are plotted upward.

4.6 Discussion

We conducted an ERP study to investigate artificial language processing. We used an AL paradigm to examine how learners acquire combinatorial constraints that govern the usage of languages and examined the influence of syntactic minimal structures using different types of adjacent dependency violations. Moreover we wanted to investigate whether abstract word order rules are extracted and are then applied through generalization, by testing implicit transfer between three-word to one and two-word

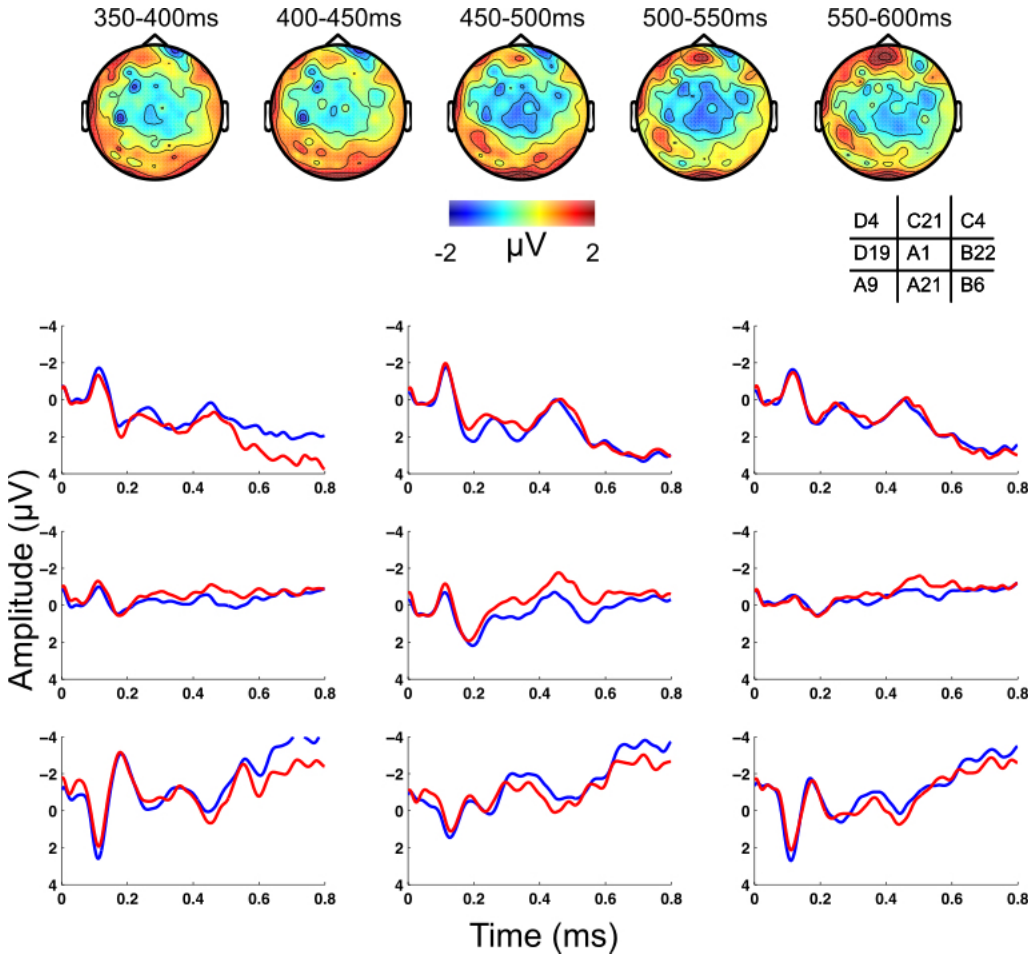


Figure 4.12: Grand-average (N=20) topographies displaying the mean amplitude difference between the ERPs evoked by the syntactic violation on the third word in three-word string in Scramble 2 and in their correct controls. Lower graphs represent grand-average (N=20) waveforms from frontal, central and parietal electrode sites time locked to the onset (0ms) third word in Scramble 2(red) and their correct controls (blue). Negative values are plotted upward.

strings.

Participants learned a miniature AL on the first day of the experiment. During training their task was to learn associations of visual stimuli and 3-pseudoword strings with a fixed order. On the following day, after approximately 24 hours of post-learning consolidation, they were tested on their knowledge: in the EEG experiment participants were presented with three types of stimuli: 1, 2 and 3-word strings. In all cases, a trial involved a visual stimulus serving as a cue for participants to recall an associated string.

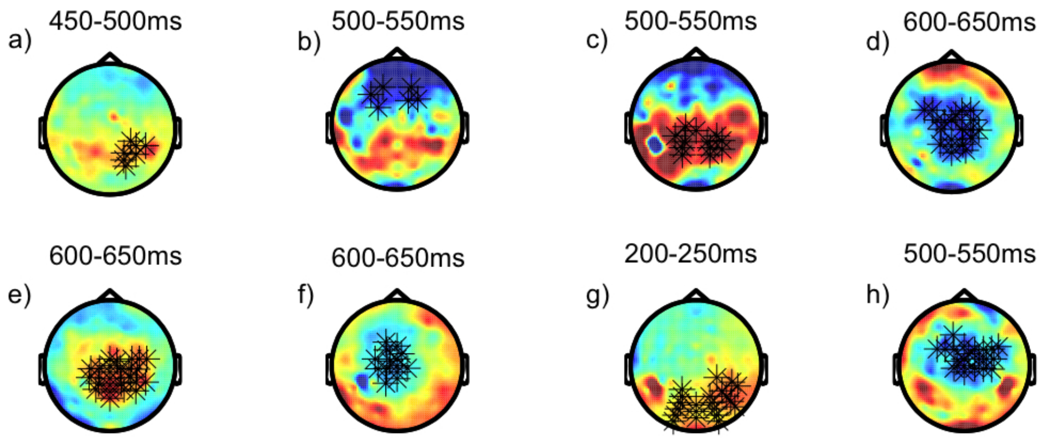


Figure 4.13: Grand-average (N=20) topographies displaying the significant clusters for (a) lexical replacement, (b) within-category replacement in one-word, (c) between-category replacement in one-word, (d) second word in Scramble in two-word strings, (e) first word in three-word strings, (f) second word in Scramble 1 in three-word, (g) second word in Scramble 2 in three-word, (h) third word in Scramble 1 in three-word.

Participants were then presented with either the correct or incorrect string and in some trials were asked to decide whether the string was correct or not. We introduced two main types of violations: lexical replacements (substitution of the correct word with a different word) and violations of word order (strings with correct lexical items, but incorrect order). Our results suggest that adults learned the implied meaning of novel words and the implied word order via a process that was sensitive to compositional constraints.

The newly acquired linguistic knowledge was integrated into their lexical knowledge. The ERP data from syntactic violations revealed that the word order knowledge was deployed rapidly in real-time sentence-like processing. The impact of the learning protocol was evident in syntactic manipulations. Following the initial exposure in the training phase to three-word strings with fixed order, a P600 effect was found in the first words in strings that exhibited violations on the first word. Interestingly, in two-word conditions we found a semantic-like effect for the incorrect word order but only on the second word of the phrase. Our interpretation is that this effect reflects knowledge of novel word meaning and order that is driven by initial exposure in the training phase to word order and the number of words in the condition.

4.6.1 Semantic manipulations

It is worth discussing our results regarding the issue of semantic manipulations. Previous studies (McLaughlin et al., 2004; Dobel et al., 2010; Bakker et al., 2015) have demonstrated that a lexical N400 effect can be obtained even in early stages of learning but leave open the question of exactly how much exposure and post-learning period is necessary for these effects to emerge. A small number of EEG studies have examined lexical learning with minimal training using AL (Borovsky et al., 2010, 2012; Mestres-Misse et al., 2007). However, in those studies only one word was novel and artificial, subjects inferred the meaning of the novel word from context of the sentence.

Previously the N400 effect was mainly found in explicit word learning tasks (Perfetti et al., 2005; Stein et al., 2006; Balass et al., 2010). Moreover, it has been suggested that the main factor contributing to the observed ERP effects relied more on episodic retrieval than lexical activation (Kiefer & Spitzer, 2000). But what about a learning environment in which all words are novel and are learned simultaneously with phrase structure through implicit learning? Our results show that semantic violations elicited an N400 effect compared to the baseline condition only for one word trials for within-category violations. The N400 effect is only evident when the semantic violation is presented without sentential context.

Two possible factors could explain this result. The first is the effect of the learning protocol. It is possible that participants, after having been exposed to the learning material composed of only three-word strings, did not regard the conflict on the third word to be strong enough. Only when the words were presented in isolation and the cognitive load was minimized was the effect present. This leads to the second possible factor: early stage of learning could partially explain this result. In our AL paradigm, participants underwent only one training on the day before the experiment which lasted around one hour.

It is difficult to compare our findings with other neurocognitive research, because previous studies have mainly used an explicit learning paradigm and, moreover, in almost all the studies the novel vocabulary was learned first and only later the syntactic structures were added. Also, not all studies consistently report the amount of training the participants received. However, from those which do, it emerges that previous AL experiments on language learning used a more extensive learning protocol extending to even several weeks (Friederici et al., 2002; McLaughlin et al., 2004). Morgan-Short et al (2012) have shown striking differences between the neural activity of artificial language learners at low and high proficiency.

For the between-category semantic manipulation we found a P600-like effect with the posterior distribution. This P600 finding proves that participants not only learned the pseudowords but also generalised the meanings to grammatical categories. The anomalous word in our experiment represented a word from a different grammatical category, thus the effect reflects a syntactic mismatch related to one word and the visual incongruous context.

Van Herten et al. (2005) proposed that a mismatch between the expected and the observed but unexpected meaning could trigger P600 effect that reflects a monitoring process that checks upon the veridicality of one's linguistic analysis. Our results coincide with Muente et al. (1997) who suggested that the functional interpretation of the P600 should be extended from a purely syntactic account to a more general reanalysis account. Several previous studies have reported violations seemingly semantic in nature that elicited P600 effect (Kim & Osterhout, 2005; Van Herten et al., 2005; Kuperberg, 2007; Schacht et al., 2014). Kuperberg (2007) suggested that P600 might reflect a linguistic processing stream that combines syntactic and semantic information in the service of comprehension. Moreover, Garnsey et al. (1989) have pointed out that processes that reflect the attempt to integrate a given grammatical category into the context are rather correlated with late positive shifts. Osterhout et al. (2004) have shown that syntactically anomalous words elicit a late positive wave. Overall this suggests that our between-category violations were perceived as syntactic in nature.

The posterior positivity effect found for the two types of lexical replacements combined together for all types of strings (1, 2- and 3-word strings) might in part reflect attentional mechanisms engaged by stimulus and task demands, which are thought to drive early positivities that represent the P300 component (Polich, 2007; Evans et al., 2011). The P300 effect has been reported in previous ERP violation studies of L2 grammar learning (Mueller et al., 2009; Morgan-Short et al., 2012). This positivity might indicate a dependence on more domain-general attentional mechanisms rather than the syntactic or semantic processing that is typical for early stage language learners. However, this interpretation may require a more detailed investigation in future studies.

The null effect we observed in our work for semantic manipulations in three-word strings is puzzling. One possibility is that the participants may have been overexposed to the correct three-word strings and they did not attend as thoroughly to the third word. This could be due to the fact that the training session consisted of only correct three-word strings and that in the EEG phase we had a higher proportion of correct trials. After seeing two correct words out of three-word strings, the participants could have processed the string more holistically, thus downplaying the manipulation of the last word. Our language had a small number of lexical items in each category. Potentially some null effects could be due to the low number of competitors leading to word category assignment that was skipped as a result of the small language size.

4.6.2 Syntactic manipulations

The finding of the positive shift effect revealed that our experiment was sensitive to manipulations in the syntactic domain but only on the first violations of the 3-word strings, as shown in Figures 4.8 and 4.10. After word order restricted training, that word order was incorporated into the grammatical systems of participant's language processing system.

One of the aims of the present study was to identify the brain patterns that underlie the processing of language-like sentences in the early stages of language learning. Previous research had shown both ELAN and late parietally distributed positivity to correlate with the processing of phrase structure information. The former reflects a highly automatic first-pass parsing (Friederici et al., 1993), while the latter is more of a controlled nature (Osterhout & Holcomb, 1992; Kaan et al., 2000). These components have been consistently found in native language speakers in response to syntactic violation (Steinhauer & Connolly, 2008).

Previous studies have shown that after an extended training these biphasic mechanisms can be found in artificial language learners (Friederici et al., 2002; Morgan-Short et al., 2012). Interestingly, this biphasic response was also found in the study with late learners of natural languages but only for learners who were exposed to the languages through immersion (Steinhauer et al., 2009). In our study, we did not find this pattern. The absence of ELAN that has been shown to reflect first syntactic parsing process is not surprising , as it has been demonstrated to be highly automatic in native speakers and rarely found in even second language learners (Hahne & Friederici, 1999; Hahne &

Friederici, 2001) and after one training our participants were still within early stages of learning. It further supports the view that the early syntactic component and late syntactic component are functionally independent (Hahne & Friederici, 2001). We found a P600 effect but no preceding anterior negativity. It may indicate that our implicit training is sufficient to establish the basic ability for structural re-analysis but it may be still at the early stage of learning under conscious control. It suggests that the amount of exposure to the new language did not result in the automatic early syntactic processing that is found in first language processing and which depends more on procedural memory.

It is worth noting that we found a P600 effect after only 1 hour of training, which is consistent with the ERP effect found usually after considerably more training of a novel natural language (4-8 months - Osterhout et al., 2008) or artificial languages (several training sessions, each session up to 5h - Friederici et al., 2002; 2 weeks - Newman-Norlund et al., 2006). One may argue that the P600 we found on the first word of the sentence is incompatible with theories implying that the effect indicates violations of phrase structure given a syntactic context. However, also Osterhout & Holcomb (1992) found P600 effects immediately after the first word, and argued that this result supports the hypothesis that any violations of preferred syntactic structures, not necessarily related to the context of the sentence, evoke a P600. It could also suggest that P600 as some have claimed reflects a more general purpose response connected with form of categorization (Schacht et al., 2014).The effect that we found only for the first word fits well with the syntactic prediction hypothesis stating that the prediction of syntactic material results from the structure that is already built in (Gibson, 1998).

A slightly delayed N400, reaching its peak around 450ms, is supported by the view that, during early stages of second language learning, syntactic processing relies partially on lexical processing and declarative knowledge (Ullman, 2004; Clahsen & Felser, 2006, Osterhout et al., 2008) and can be a result of implicit learning (Morgan-Short et al. 2012). We observed an N400 for the string-final words in Scramble 1 and 2, and a negative shift in the second word in two-word strings for Scramble. It has been suggested that, at the early stage of learning, syntactic violations elicit N400 effects. The structural violation at the early stage of learning is not yet recognized as such and is processed as lexical violation (Steinhauer et al., 2009). As suggested by Ullman (2005), N400s should be the standard response to any syntactic anomaly in early stages of learning. Only later, the beginning of grammaticalization and proceduralization starts, learners classify the violation as structural,and possibly attempt to repair the problem. The P600 will then reflect one (or several) of these emerging processes.

The syntactic anomaly at the early stage of language learning is recognized as a lexical one and declarative rule knowledge is responsible, as procedural memory is not yet accessible (Osterhout et al., 2006; Steinhauer et al., 2009). Opitz & Friederici (2003) found in their fMRI study that during the early stages of artificial language learning participants showed initial activation in declarative memory structures and the activation in procedural memory increased only after more training. As our participants were in the early stage of novel language learning, this may provide further support for the view that AL (as well as second language) learning, at least in these early stages, relies on lexical processing and declarative memory. Similarly, Weber-Fox & Neville (1996) found no ELAN and P600 but a significant negativity between 300-500 ms for late bilingual non-native speakers.

Kim and Osterhout (2005) have concluded that under certain conditions semantic information is in control of how words are combined during syntactic processing. Our results suggest that the early stage of learning could be one of them. Hahne and Friederici (2001) suggested that in the early learning stage, the learners may be activating additional processes on the basis of conceptual processing. It also has been shown that at higher proficiency level this processing may come to depend more on native-like syntactic processing and procedural memory (Friederici et al., 2002; Ullman, 2004; Steinhauer et al., 2009; Morgan-Short et al., 2012).

We found the N400 effect for the third words in two Scramble conditions, as opposed to two Displacement conditions in which this effect was not present. This could suggest that, only when the violation is more incongruous (two dependencies are violated instead of one), the ERP effects between correct and incorrect sentences can be observed. It is possible that the violations in Displacement conditions were not salient enough to trigger the neurocognitive processes that are reflected by the ERP effect. Moreover, some authors have suggested that language-related ERP effects are less likely to be elicited in language learners when the grammatical violations of a new language are difficult to transfer from the L1 of the learners (Mueller, 2005; Vaughan-Evans et al., 2014; Caffarra et al., 2015) as could be a case in Displacement conditions.

The negative effect in the two-word condition elicited on the second word lasted slightly longer (220 - 798 ms) than reported by earlier studies, however the longer extension of the N400 was also found by Kutas & Kluender (1991) and Hahne & Friederici (2001) and was suggested to be the result of uncertainty. This long extension may indicate that learners need more time to integrate the word into its prior context when the presented string (2-word long) was not included in the training material.

Together, the results from syntactic manipulations suggest that the first violations in the syntactic structure elicit a P600 effect. Subsequent violations lead to N400 effects and a reliance of grammar on lexical processing. Secondly, the syntactic manipulations of the strings that were not present in the training material (2-word strings) show the negative long lasting shift, similarly showing the dependence on lexical processing. Overall, the results suggest that brain processing of syntactic violations of sentential order depends on the position of the violation as well as the training material.

4.6.3 Future directions and improvements

Linguistic knowledge is consolidated through years of experience. Our experiment was composed of merely two sessions, none of which involved language production. Caffarra et al. (2015) suggested that proficiency and immersion are important factors in language learning. Our paradigm was a very simplified method for testing the early stages of language learning. In a natural learning context, the lexical knowledge is often incrementally reinforced after multiple exposures and usage in a variety of contexts (Nagy et al., 1985). Even though our words were presented in different contexts it is possible that the context was not altered enough to fully enrich the understanding of the words. To solve one of the issues, one could train the participants in an extended training via signaling games, where they would undergo the training as both the receiver (as in current study) and as the sender. This would give the participants the chance to boost their new knowledge not only from mere exposure but also through usage.

One of our methodological concerns is the monitoring control task, in which subjects were required from time to time answer the question about the correctness of the stimulus items. We used this task to monitor attention but this may focus the participant's attention on the sentence's form and meaning in way that is different in the communicative use of language. In future studies, control tasks focusing on the language itself could allow a more ecologically valid testing environment.

Some of our null effects could be the result of tiredness and learning strategies not intended by us but as a side effect of the long duration of the EEG experiment. Another potential concern is the visual presentation in sentence-like processing. To avoid eye movements, the sentences were presented word after word with the presentation slower than in normal reading (300ms). This word-by-word presentation could have induced a higher load on memory than what would be present natural language. However, it is difficult to see how this would induce a change in the brain response to the violations.

We had approximately a 24h post-learning consolidation period between the training session and EEG phase. It is possible that this was not enough to fully assimilate the novel language in the neocortical systems. Kaczer et al. (2015) had a longer (48h) period and still found differences between novel and familiar words and concluded that the consolidation of new knowledge was incomplete. It would be interesting to investigate what is the appropriate amount of consolidation period for artificial languages.

The current study focused on syntactic and semantic processing of a newly acquired miniature artificial language. Even though we tried to minimize the influence of confounding factors, (for example world knowledge, exposure or frequency) other factors could have influenced the learning and processing in our experiment. One possible explanation of some of our null effects could be potential individual differences in language processing and its reflection in the electrophysiological responses. As a common practice, averaged ERP waveforms are the result of averaging over participants. However, some studies have demonstrated that groups of participants may vary in the way individuals process language (Osterhout, 1997; Nieuwland & Van Berkum, 2005) and that ERPs appear to be sensitive to these differences (Osterhout, 1997). Depending on the linguistic manipulations, these effects have been connected with a range of various aspects, such as memory, proficiency, cognitive control, cognitive load and even genetic effects (Fisher et al., 2003; Kolk et al., 2003; Snijders et al., 2009). It is possible that the inter-individual differences are even more robust in artificial language studies, where participants could adopt various cognitive strategies in learning a new language and then solving the experimental task. It is an important factor that should be taken into account in future studies investigating artificial language learning and processing in order to be truly able to generalize results to natural languages.

Overall, our results are in agreement with previous studies on artificial and second language learning that suggest the lexical processing is similar to L1 processing, but the neural processes underlying syntactic could be highly dependent on the learning material and type of exposure. Further studies have to examine the amount of training and exposure necessary to achieving native-like neural effects. Participants learned a novel language, they showed a high accuracy in the control questions, demonstrating the not only had they learned the lexical content but also the implied word order. Our paradigm appears to be a good method to test for implicit language learning; learners had to extract the syntactic structure from strings they were exposed to and map them to semantic structure as well as generalise them to 2-word strings. In addition, it highlights the importance of examining the early stages of language learning to better understand the critical issues of basic linguistic processing. The use of ERPs and the AL paradigm in the current study confirms the importance of EEG in studying language learning in adults.

In summary, we demonstrated the existence of a rapid recruitment of language-related brain networks for the processing of novel miniature languages. We found an effect that relates to the semantic learning of novel words that was found in a one-word condition. ERP results show the sensitivity to syntactic structures, with certain limitations, highlighting the importance of implicit learning and learning material. We believe that research on language learning and processing should be conducted in an ecologically valid way. This sort of implicit learning method with additional elements of communicative pressure (in the form of signaling games training) could in the future reveal the complex interactions between the different factors involved in adult language learning.

Chapter 5

Discussion

5.1 Summary of the results

Taken together, the results reported in this thesis suggest that artificial language learning can shed light on the cognitive and social factors that constrain language learning and language change. Although evidence for this conclusion come from various experimental sources, we have shown that artificial language learning studies provide a way to test hypotheses about constraints responsible for shaping natural languages. Our work targeted fundamental and universal properties of natural languages such as constituent order and compositionality. These are prime test cases for exploring the existence of learning biases.

In the first study (Chapter 2) our main goal was to assess the role of learning constraints favoring structural order, or disfavoring linear order. The widely accepted assumption is that grammatical operations are generally structure-dependent and do not exploit linear dependencies. We used a children-friendly puppet theatre and the task was adjusted to the children's age. We found that children are not unconstrained learners. Children showed a preference for strings whose structure they were exposed to during training phase, but this preference was stronger for structural and free order grammars than for linear order grammars. The effect was modulated by the children's first language. Our results indicate that structural and free order grammars may be easier to learn than linear order grammars, possibly indicating the existence of constraints favouring the typologically plausible constituent order. Interestingly, this effect was not found in Italian-, Polish- or German- speaking adults.

In the second study (Chapter 3) we set out by investigating whether certain linguistic universals can be studied experimentally using models of code transmission in diffusion chains. We introduced a novel experimental paradigm, the iterated signaling game, to investigate how simple artificial languages emerge and change in the course of horizontal and vertical transmission. We showed that, by adding communicative pressure and feedback to iterated learning, the emerging codes were highly structured and expressive right from the first generation. In the first experiment we showed an increase in coordination and transmission: the codes further down the chain were becoming easier to learn and transmit. In the second experiment, where we introduced basic grammatical functions, we found the emergence of morphological marking and a decrease of innovation throughout the chains. The emergence of proto-morphology suggests that one of its functions may be the facilitation of communication: participants throughout a chain restructured the code and used word structure to aid evolution towards a shared communicative system. Overall, we provided a proof of concept for iterated signaling games as an experimental model of language change, and we established a baseline of results for further research employing iterated signaling games.

The last study (Chapter 4) extended previous results by exploring brain responses in the initial stages of learning an artificial language. The learning paradigm we employed made further use of signaling games. We conducted an artificial language learning experiment on two days to examine the neurophysiological foundations of basic compositional processing of artificial language-like structures. Our results show that the learning of grammatical regularities can be fast: in a single training session, participants learnt semantic content and extracted the governing syntactic rule. Moreover, our results are in accordance with previous studies on artificial language learning, suggesting that lexical processing can be similar to L1 processing in the initial stages of non-native language learning, resulting in native-like, albeit slightly delayed (as predicted by previous research), N400 effects. Syntactic processing, however, can be influenced by the type of material the learners are exposed to during the learning stage. The ERP data from syntactic manipulations revealed that syntactic knowledge was deployed rapidly, yielding P600 effects being highly affected by the learning protocol.

5.2 Theoretical implications

The results we obtained in Chapter 2 are consistent with the existence of developmental constraints on learning favouring structural constituent order. As a secondary research question one might ask what are the factors that influence learning. Our results point to two main answers: (1) the age of participants - children but not adults showed preference in learning structural order rules - and (2) the learners' L1. Moreover, it is worth mentioning that semantic cues were not necessary in order to learn governing rules when anchor points, such as function and content words, were available.

Chapter 3 investigated language change using iterated signaling games. One of the most interesting results we obtained was the emergence of proto-morphology. Morphology, in languages exhibiting a case marking system, is the main strategy languages employ to disambiguate between object and subject by explicit markings. Communicative pressures have been proposed to explain how case marking systems arise (Comrie, 1989; Jager, 2007). Our results fit with these theories and furthermore suggest that the 'morphologization' process contributes only when there is a need to reduce ambiguity in communication.

We also showed that rapid changes in vocabulary occur during horizontal transmission, while properties of language that are necessary for communication emerge quickly and remain stable. We believe that the results of experiments described in Chapter 3 have a broad scope and could be relevant for the study of all forms of communication, as they can contribute to the understanding of the key properties of natural languages. The novel codes that emerged in our experiments after introducing communicative pressure share important similarities with natural languages. In addition, our results replicate earlier findings by Moreno & Baggio (2014), providing further evidence that language change in signaling games is mostly unidirectional from senders to receivers while extending those findings to diffusion chains.

The experiment described in Chapter 4 provides evidence for the power of language learning mechanisms. Our ERP results suggest that non-native language learning is partially dependent on the learning material and it seems to be dependent also on lexical processing during early stages of novel language learning. This is in line with earlier theories (Ullman, 2004). A positive result is that participants, after just one short learning session through signaling games, can exhibit ERP effects found in the

4th stage out of 6 stages of an L2 learning span proposed by Steinhauer et al. (2009).

5.3 Methodological implications

Our results underscore the use of the artificial languages as a valuable source of empirical data for language learning. One of our aims in the first study (Chapter 2) was to design artificial languages that were somehow closer to natural languages but also simple enough to be learnable in a training session of reasonable length and to be suitable to preschool children. Our paradigm introduced key properties of natural languages (the function vs content words distinction) and takes artificial languages a step closer to natural languages while maintaining a high degree of control.

The experiments reported in Chapter 3 have important methodological implications. They show the importance of communicative pressure in language transmission experiments. We provided evidence that iterated signaling games could serve as a viable experimental model of language change within a collaborative environment. Iterated signaling games have important properties (flexibility, simplicity, formal explicitness) and allows us to test how learners update their knowledge and how the interactions influence it.

Up until now most experiments on cultural transmission were conducted using only behavioral measures. Our EEG study was a first small step towards justifying signaling games as a valid technique to study neural correlates of cognitive activities involved in language transmission. Chapter 4 underlines the importance of combining EEG and artificial languages. The ERP results suggest that our participants were in the early stages of learning the language. Importantly, our study also shows that artificial language learning taps into similar cognitive mechanisms as those involved in natural language learning. Furthermore, it shows that signaling games are an effective paradigm to test language learning: participants learnt the rule encountered in the training phase (3-word strings) quickly and also transferred the rule to 2-word strings effectively. Moreover, our more naturalistic learning of both semantic content and syntactic structures simultaneously suggests that it is possible to minimize the length of training in artificial learning to progress in learning the language.

5.4 Future outlook

Here I will briefly describe certain limitations and possible future directions for our described experiments and methods.

In the first study (Chapter 2), while we are sure the participants have learned some aspects of the artificial language, it is difficult to characterize the nature of knowledge that results from the artificial language learning. The difficulty is due to several reasons which we will now describe.

The aim was to provide evidence for the existence of learning constraints. However, the results do not allow me to draw conclusions about their nature or their origin. Nevertheless, there are two possible hypotheses. One is that learners had abstract knowledge of the rule. They have successfully generalized to a novel vocabulary, suggesting that they have acquired a surface-independent representation of the rule. Moreover, the strings we used were from two to five words long which allowed for even more abstract generalization.

The second hypothesis is that participants acquired the rules based on 'micro-rules'. Micro-rules are rules based on the perceptual characteristics of the stimuli that only partially account for the sentences generated by the target rule (Dulany et al., 1984; Pothos, 2007). It suggests that participants may have acquired only some aspects of the target rule. For instance in our experiments the learning strategy that participants could have potentially adopted was to learn for example that the linear order starts with a function word. However, these micro-rules cannot explain the differences in learning between two different orders or the imperfect performance of the participants.

The performance was rarely perfect (few participants had 100% correct responses). It is possible that some of our subjects had full competence of the rule but that there were cognitive limitations influencing their performance. These limitations may include attention, the quality of the information, memory limitations or even the lack of consolidation period between training and test phases (Friedrich et al., 2015). There was a series of methodological choices that had to be made which may have influenced some of our results. Most studies in artificial grammar learning use passive training in which participants are simply asked to listen to lists of strings generated by the grammars. This method obviously has some shortcomings: it is far removed from naturalistic language learning, where learners are immersed in a meaningful context; it may be a challenge for children to maintain attention throughout the training phase in passive learning environments.

From a more practical perspective, the test phase could be improved by increasing the number of trials to facilitate the interpretation of the learning outcome. Due to the age of our participants, their short attention span and the presence of feedback, we had to exclude that option. Furthermore, we did not have semantic cues, even though function and content words could have partially replaced their role. The semantic bootstrapping hypothesis (Pinker, 1984) proposes that, during language learning, learners may rely on the correspondence between semantic and syntactic information to learn, for example, syntactic categories. So far not many studies investigated the role of semantics in artificial language learning and reached contradictory conclusions. In our experiment we opted not to use semantic content, the next step should be to introduce semantic cues to potentially facilitate the learning of the rules.

An obvious extension would be to investigate the nature of generalization and abstraction that participants make (lexical generalization, syntactic generalization or both). Clearly, more data are needed to further explore the learnability of constituent orders and to characterize the processing of these structures. Neuroimaging studies could explore how learning different constituent orders could generate changes in brain activity (Tettamanti et al., 2002).

The participants in Chapter 3 and 4 were adults with a life-long experience of their native language and social communication. This may have impacted the final outcome of the results. Most of the studies within the artificial languages framework have been conducted with adults. This trend has been changing to include children, infants and even newborns (Saffran et al., 1996; Gomez & Gerken, 1999; Ferry et al., 2015) which allows one to make comparison about the language learning mechanisms and its developmental trajectory. Despite being scientifically challenging, further studies could explore the language transmission by children. There are no studies (to our knowledge) that directly tested children's performance and ability to learn languages via cultural transmission.

In experiments described in Chapter 3 we investigated the simplest type of iterated signaling games: each generation consisted of only two individuals with fixed roles. There are several interesting extensions that could be easily implemented with iterated signaling games. Firstly, one may extend signaling games to a larger number of players into signaling networks and explore how the final output changes (Skyrms 2010). Secondly, one could use more complex syntactic structures and larger languages to investigate the boundaries of language learning through transmission. Thirdly, by manipulating the frequency of the input condition, it may be possible to experimentally explore the extent to which differences in the frequency of some inputs affect the change, as shown by Pagel et al. (2007). Finally, it would be important to establish the relative contributions of interactions and feedback. Until now these two aspects were usually studied together but it is possible that one or the other is driving the results. Overall, a key challenge for the future is to further isolate the specific factors that contribute to language transmission.

In follow-up studies presented in Chapter 4, certain methodological issues could be improved to make sure the learning process can be monitored. Our results suggest that participants were within the early stage of language learning at the end of our training, future studies would require more extensive training. In addition to more studies combining artificial language learning and brain imaging techniques, more longitudinal electrophysiological studies, tracking learners' progress through different levels of proficiency, are needed. It would help to establish the stages of learning and the methodological endpoint of training when participants could be tested and the results generalised to native-language speakers. Artificial language studies which by design ensure a high learning proficiency (Friederici et al., 2002; Morgan-Smith et al., 2012) is a promising area of research.

Miniature artificial languages are an important experimental tool as they limit the amount of training to minimum, compared with second language learning, while still controlling the amount and type of exposure. The time-frame of reaching native-like processing and proficiency becomes feasible. Our learning paradigm seems to shorten the amount of exposure needed to elicit syntactic ERPs (however one has to remember that our language was relatively small and simple). A clear and obvious extension to our experimental setup would involve adding a social component in the training phase to investigate whether the learning protocol (social vs nonsocial) determines brain electric response or speeds up the language learning.

Although every experimental approach has its strength and weaknesses, we believe that artificial languages used in conjunction with neuroimaging techniques, provide a rare opportunity to make progress in understanding the neurocognitive bases of language learning.

References

Amato, M. S., & MacDonald, M. C. (2010). Sentence processing in an artificial language: Learning and using combinatorial constraints. *Cognition*, 116(1), 143–148.

Ambridge, B., Rowland, C. F., & Pine, J. M. (2008). Is structure dependence an innate constraint? New experimental evidence from children's complex-question production. *Cognitive Science*, 32, 222-255.

Amoruso, L., Gelormini, C., Aboitiz, F., Alvarez Gonzalez, M., Manes, F., Cardona, J. F.. & Ibanez, A. (2013). N400 ERPs for actions: building meaning in context. *Frontiers in Human Neuroscience*, 7(57), 1-16.

Arbib, M. A. (2005). *From monkey-like action recognition to human language: An acquisition and change.* Amsterdam: Benjamins.

Ardal, S., Donald, M.W., Meuter, R., Muldrew, S. & Luce, M. (1990). Brain responses to semantic incongruity in bilinguals. *Brain and Language*, 39, 187–205.

Atkinson, Q. D., Meade, A., Venditti, C., Greenhill, S. J., & Pagel, M. (2008). Languages evolve in punctuational bursts. *Science*, 319(5863), 588.

Bahlmann, J., Schubotz, R. I., & Friederici, A. D. (2008). Hierarchical artificial grammar processing engages Broca's area. *NeuroImage*, 42(2), 525–534.

Bakker, I., Takashima, A., van Hell, J. G., Janzen, G., & McQueen, J. M. (2015). Tracking lexical consolidation with ERPs: Lexical and semantic priming effects on N400 and LPC responses to newly-learned words. *Neuropsychologia*, 79, 33–41.

Balass, M., Nelson, J.R., & Perfetti, C.A. (2010). Word learning: an ERP investigation of word experience effects on recognition and word processing. *Contemporary Educational Psychology*, 35 (2), 126–140.

Baronchelli, A., Chater, N., Pastor-Satorras, R., & Christiansen, M. H. (2012). The Biological Origin of Linguistic Diversity. *PLoS ONE*, 7(10).

Barr, D.J. (2004). Establishing conventional communication systems: Is common knowledge necessary? *Cognitive Science*, 28, 937–962.

Bartlett, F.C. (1932). *Remembering: A Study in Experimental and Social Psychology.* Cambridge: Cambridge University Press.

Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., & Schoenemann, T. (2009). Language is a complex adaptive system: Position paper. *Language Learning*, 59(SUPPL. 1), 1–26.

Beckner, C., Rcz, P., Hay, J., Brandstetter, J. & Bartneck, C. (2015). Participants conform to humans but not to humanoid robots in an English past tense formation task. *Journal of Language and Social Psychology*, 8, 1-22.

Bell, A., Brenier, J. M., Gregory, M., Girand, C., Jurafsky, D. (2009). Predictability effects of duration of content and function words in conversational English. *Journal of Memory and Language*, 60, 92-111.

Bemis, D. K., & Pylkknen, L. (2011). Simple composition: a magnetoencephalography investigation into the comprehension of minimal linguistic phrases. *The Journal of Neuroscience*, 31(8), 2801–2814.

Berman, R. A., & Slobin, D. (1994). *Relating events in narrative: a crosslinguistic developmental studies.* Hillsdale, NJ: Lawrence Earlbaum Associates.

Bickerton, D. (1990). *Language and Species.* Chicago: University of Chicago Press.

Bickerton, D. (2003). Symbol and structure: A comprehensive framework for language evolution. In M. H. Christiansen & S. Kirby (Eds.), *Language evolution* (pp. 77–93). Oxford: Oxford University Press.

Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language*, 68, 706-755.

Borovsky, A., Elman, J. L., & Kutas, M. (2012). Once is enough: N400 Indexes Semantic Integration of Novel Word Meanings from a Single Exposure in Context. *Language Learning Development*, 8(3), 278–302.

Borovsky, A., Kutas, M., & Elman, J. (2010). Learning to use words?: Event-related potentials index single-shot contextual word learning. *Cognition*, 116(2), 289–296.

Bowie, J. (2008). Proto-discourse and the emergence of compositionality. *Interaction Studies*, 9(1), 18-33.

Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, 108 Suppl , 10918–10925.

Braine, M. D. S. (1966). Learning the position of words relative to a marker element. *Journal of Experimental Psychology*, 72(4), 532-540.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, 75, 13–25.

Brennan, S. E., & Hanna, J. E. (2009). Partner-specific adaptation in dialogue. *Topics in Cognitive Science* 1, 274–291.

Bybee, J. (2007). *Frequency of use and the organization of language.* Oxford: Oxford University Press.

Bybee, J. (2012). Domain-general processes as the basis for grammar. In: M. Tallerman & K. Gibson (Eds.), *The Oxford Handbook of Language Evolution* (pp. 528-536). Oxford: Oxford University Press.

Caffarra, S., Molinaro, N., Davidson, D., & Carreiras, M. (2015). Second language syntactic processing revealed through event-related potentials: An empirical review. *Neuroscience & Biobehavioral Reviews*, 51, 31–47.

Caldwell, C., & Millen, A. E. (2008). Review. Studying cumulative cultural evolution in the laboratory. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 363(1509), 3529–3539.

Carrington, S. J., & Bailey, A. J. (2009). Are there theory of mind regions in the brain? A review of the neuroimaging literature. *Human Brain Mapping*, 30(8), 2313–2335.

Chapman, R. S., & Kohn, L. L. (1978). Comprehension Strategies in Two and Three Year Olds: Animate Agents or Probable Events? *Journal of Speech, Language and Hearing Research*, 21, 746-761.

Chater, N., & Christiansen, M. H. (2010). Language Acquisition Meets Language Evolution. *Cognitive Science*, 34(7), 1131–1157.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Christiansen, M. H. (2000). Using artificial language learning to study language evolution: Exploring the emergence of word order universals. In Dessalles, J. & Ghadakpour, L. (Eds.), *The Evolution of Language.* Paris: Ecole Nationale Superieure des Telecommunications.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *The Behavioral and Brain Sciences*, 31(5), 489–508; discussion 509–558.

Clahsen, H., & Felser, C. (2006). How native-like is non-native language processing? *Trends in Cognitive Sciences*, 10(12), 564–570.

Clark, B., Goldrick, M., & Konopka, K. (2008). Language Change as a Source of Word Order Correlations. *Variation, Selection, Development: Probing the Evolutionary Model of Language Change*, 75–102.

Clark, H. H. (1996). *Using Language.* Cambridge: Cambridge University Press.

Clark, H.H., Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1–39.

Comrie, B. (1981). *Language universals and linguistic typology.* Chicago: University of Chicago Press.

Cooper, D. (1999). *Linguistic attractors: The cognitive dynamics of language acquisition and change.* Amsterdam, Netherlands: John Benjamins.

Cornish, H., Tamariz, M., & Kirby, S. (2009). Complex adaptive systems and the origins of adaptive structure: What experiments can tell us. *Language Learning*, 59(SUPPL. 1), 187–205.

Costa, A., & Sebastian-Galles, N. (2014). How does the bilingual experience sculpt the brain? *Nature Reviews Neuroscience*, 15, 336-345.

Crain S. (1991). Language acquisition in the absence of experience. *Behavioral and Brain Sciences*, 4, 597–650.

Crain, S., & Nakayama, M. (1987). Structure dependence in grammar formation. *Language*, 63, 522-543.

Crain, S., & Thornton, R. (2000). *Investigations in Universal Grammar: A guide to Experiments on the Acquisition of Syntax.* Cambridge, MA: MIT Press.

Croft, W. (2009). Towards a social cognitive linguistics. In V. Evans & S. Pourcel (Eds.), *New directions in cognitive linguistics* (pp. 395–420). Amsterdam: Benjamins

Culbertson J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition*, 122(3), 306-329.

Culbertson, J. (2012). Typological Universals as Reflections of Biased Learning: Evidence from Artificial Language Learning. *Linguistics and Language Compass*, 6(5), 310–329.

Cuskley, C. F., Pugliese, M., Castellano, C., Colaiori, F., Loreto, V., & Tria, F. (2014). Internal and external dynamics in language: Evidence from verb regularity in a historical corpus of english. *PLoS ONE*, 9(8).

Dale, R. & Christiansen, M.H. (2004). Active and passive statistical learning: Exploring the role of feedback in artificial grammar learning and language. In *Proceedings of the 26th Annual Meeting of the Cognitive Science Society* (pp. 262-267). Mahwah, NJ: Lawrence Erlbaum.

De Ruiter, J. P., Noordzij, M. L., Newman-Norlund, S., Newman- Norlund, R., Hagoort, P., & Levinson, S. C. (2010). Exploring the cognitive infrastructure of communication. *Interaction Studies*, 11(1), 51-77.

de Vries, M. H., Monaghan, P., Knecht, S., & Zwitserlood, P. (2008). Syntactic structure and artificial grammar learning: The learnability of embedded hierarchical structures. *Cognition*, 107(2), 763-774.

Dediu, D., & Levinson, S. C. (2012). Abstract Profiles of Structural Stability Point to Universal Tendencies, Family-Specific Factors, and Ancient Connections between Languages. *PLoS ONE*, 7(9).

Dekeyser, R. M. (2000). The Robustness of Critical Period Effects in Second Language Acquisition. *Studies in Second Language Acquisition*, 22, 499–533.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015a). Arbitrariness, Iconicity, and Systematicity in Language. *Trends in Cognitive Sciences*, 19(10), 603–615.

Dingemanse, M., Roberts, S. G., Baranova, J., Blythe, J., Drew, P., Floyd, S., & Enfield, N. J. (2015b). Universal Principles in the Repair of Communication Problems. *PLoS ONE*, 10(9).

Dobel, C., Junghofer, M., Breitenstein, C., Klauke, B., Knecht, B., Pantev, C., & Zwitserlood, P. (2010). New names for known things: on the association of novel word forms with existing semantic information. *Journal of Cognitive Neuroscience*, 22(6), 1251-1261.

Donchin, E. (1981). Surprise...surprise? *Psychophysiology*, 18, 493-513.

Dryer, M. (1992). The Greenbergian word order correlations. *Language*, 68, 81-183.

Dulany, D. E., Carlson, R. A., & Dewey, G. I. (1984). A case of syntactical learning and judgement: How conscious and how abstract? *Journal of Experimental Psychology: General*, 113, 541-555.

Dunn, M., Greenhill, S. J., Levinson, S. C., & Gray, R. D. (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, 473(7345), 79–82.

Ellis, N. C., & Cadierno, T. (2009). Constructing a Second Language. *Annual Review of Cognitive Linguistics*, 7(1), 111–139.

Enrici, I., Adenzato., M., Cappa, S., Bara, B., G., & Tettamanti, M. (2011). Intention processing in communication: a common brain network for language and gestures. *Journal of Cognitive Neuroscience*, 23(9), 2415-2431.

Evans, J. L., Selinger, C., & Pollak, S. D. (2011). P300 as a measure of processing capacity in auditory and visual domains in specific language impairment. *Brain Research*, 1389, 93–102.

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, 329, 429-492.

Fava, E., Hull, R., & Bortfeld, H. (2011). Linking behavioral and neuropsychological indicators of perceptual tuning to language. *Frontiers in Psychology*, 2(AUG), 1-14.

Fay, N., & Ellison, T. M. (2013). The Cultural Evolution of Human Communication Systems in Different Sized Populations: Usability Trumps Learnability. *PLoS ONE*, 8(8).

Fay, N., Garrod, S., & Carletta, J. (2000). Group discussion as interactive dialogue or as serial monologue: the influence of group size. *Psychological Science*, 11(6), 481–486.

Fay, N., Garrod, S., & Roberts, L. (2008). The fitness and functionality of culturally evolved communication systems. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 363(1509), 3553–3561.

Fay, N., Garrod, S., Roberts, L., & Swoboda, N. (2010). The interactive evolution of human communication systems. *Cognitive Science*, 34(3), 351–386.

Federmeier, K. D., & Kutas, M. (1999). A Rose by Any Other Name: Long-Term Memory Structure and Sentence Processing. *Journal of Memory and Language*, 495, 469–495.

Federmeier, K. D., & Kutas, M. (2001). Meaning and Modality: Influences of Context, Semantic Memory Organization, and Perceptual Predictability on Picture Processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 27(1), 202-224.

Fedor, A., Varga, M., & Szathmary, E. (2012). Semantics boosts syntax in artificial grammar learning tasks with recursion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(3), 776-782.

Fedzechkina, M., Jaeger, T. F., & Newport, E. L. (2012). Language learners restructure their input to facilitate efficient communication. *Proceedings of the National Academy of Sciences*, 109(44), 17897–17902.

Fehr, O., Wang, H., Saar, S., Mitra, P. P., & Tchernichovski, O. (2009). De novo establishment of wild-type song culture in the zebra finch. *Nature*, 459(7246), 564–568.

Ferry, A. L., Flo, A., Brusini, P., Cattarossi, L., Macagno, F., Nespor, M., & Mehler, J. (2015). On the edge of language acquisition: inherent constraints on encoding multisyllabic sequences in the neonate brain. *Developmental Science*, 6, 1-16.

Finley, S., & Badecker, W. (2009). Artificial language learning and feature-based generalization. *Journal of Memory and Language*, 61, 3, 423-437.

Finn, A. S., Lee, T., Kraus, A., & Hudson Kam, C. (2014). When It Hurts (and Helps) to Try: The role of Effort in Language Learning. *PLoS ONE*, 9(7).

Fitch, W. T. (2007). Linguistics: An invisible hand. *Nature*, 449(7163), 665–667.

Fletcher, P. C., Zafiris, O., Frith, C., D., Honey, R. A., Corlett, P., R., Ziler, K., & Fink, G., R. (2013). On the Benefits of not Trying: Brain Activity and Connectivity Reflecting the Interactions of Explicit and Implicit Sequence Learning. *Cerebral Cortex*, 15(7).

Flynn, E., & Whiten, A. (2010). Studying children's social learning experimentally 'in the wild'. *Learning & Behavior: A Psychonomic Society Publication*, 38(3), 284–296.

Folia, V., Udden, J., De Vries, M., Forkstam, C., & Petersson, K. M. (2010). Artificial language learning in adults and children. *Language Learning*, 60, 188-220.

Frank, M. C., & Gibson, E. (2011). Overcoming Memory Limitations in Rule Learning. *Language Learning and Development,* 7(2), 130-148.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6(2), 78–84.

Friederici, A. D. (2005). Neuropsychological markers of early language acquisition: from syllables to sentences. *Trends in Cognitive Science*, 9, 481-488.

Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: Early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1219–1248.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. Brain Research. *Cognitive Brain Research*, 1(3), 183–192.

Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences*, 99(1), 529–534.

Friedrich, M., Wilhelm, I., Born, J., & Friederici, A. D. (2015). Generalization of word meanings during infant sleep. *Nature Communications*, 6, 6004.

Futrell, R., Mahowald, K., & Gibson, E. (2015). Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences*, 112, 10336-10341.

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science*, 29(5), 737–767.

Galantucci, B., & Garrod, S. (2011). Experimental semiotics: a review. *Frontiers in Human Neuroscience*, 5,11.

Galantucci, B., Garrod, S., & Roberts, G. (2012). Experimental Semiotics. *Language and Linguistics Compass*, 6(8), 477-493.

Garnsey, S. M.,Tanenhaus, M. K., & Chapman, R. M. (1989).Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research*, 18, 51-60.

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: a study in conceptual and semantic coordination. *Cognition*, 27(2), 181–218.

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, 8(1), 8–11.

Garrod, S., Fay, N., Lee, J., Oberlander, J. & Macleod, T. (2007). Foundations of representation: where might graphical symbol systems come from? *Cognitive Science*, 31, 961–988.

Garrod, S., Fay, N., Rogers, S., Walker, B., & Swoboda, N. (2010). Can iterated learning explain the emergence of graphical symbols? *Interaction Studies*, 11(1), 33-50.

Gervain, J., Sebastian-Galles, N., Diaz, B., Laka, I., Mazuka, R., Yamane, N., Nespor, M., & Mehler, J. (2013). Word frequency cues word order in adults: cross-linguistic evidence. *Frontiers in Psychology*, 4, 689.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68, 1–76.

Giles, H., Coupland, N. & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. *Contexts of accommodation*, 28(2), 105 – 24.

Gomez, R. L. & Gerken, L. A. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, 4, 178-186.

Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, 13(5), 431-436.

Gomez, R. L., & Gerken, L. A. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, 70, 109-135.

Gomez, R. L., Gerken, L., & Schvaneveldt, R. W. (2000). The basis of transfer in artificial grammar learning. *Memory & Cognition*, 28, 253-263.

Gouvea, A. C., Phillips, C., Kazanina, N., & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes*, 25(2), 149–188.

Green, T. R. G. (1979). Necessity of syntax markers: Two experiments with artificial languages. *Journal of Verbal Learning and Verbal Behavior*, 18, 481-496.

Greenberg, J. (1959). Current trends in linguistics. *Science*, 130, 1165-1170.

Greenberg, J. (1963). *Universals of Language.* Cambridge: Cambridge University Press.

Griffiths, T. L., Kalish, M. L., & Lewandowsky, S. (2008). Review. Theoretical and empirical evidence for the impact of inductive biases on cultural evolution. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 363(1509), 3503–3514.

Grimshaw, J. (1997). Projection, heads and optimality. *Linguistic Inquiry*, 28, 373-422.

Hagoort, P. (2003). How the brain solves the binding problem of a language: a neurocomputational model of syntactic processing. *NeuroImage*, 20(S1): S18-S29.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, 8(4), 439–483.

Hagoort, P., Wassenaar, M., & Brown, C. M. (2003). Syntax-related ERP-effects in Dutch. *Cognitive Brain Research*, 16(1), 38–50.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, 11(2), 194–205.

Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, 4(02), 123–141.

Healey, P. G. T., Garrod, S., Fay, N., Lee, J., & Oberlander, J. (2002). Interactional context in graphical communication. *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (pp. 441– 446). Mahwah, NJ: Lawrence Erlbaum Associates.

Healey, P. G. T., Swoboda, N., Umata, I., & King, J. (2007). Graphical language games: interactional constraints on representational form. *Cognitive Science*, 31, 285–309.

Heine, B. & Kuteva, T. (2002). *Word lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Heine, B., & Kuteva, T. (2007). *The Genesis of Grammar: A Reconstruction*. Oxford: Oxford University Press.

Heine, B., Claudi, U., & Hunnemeyer, F. (1991). *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.

Hochmann, J. R., Endress, A. D., & Mehler, J. (2010). Word frequency as a cue for identifying function words in infancy. *Cognition*, 115(3), 444-457.

Hockett, C. (1960). *The origin of speech*. Scientific American, 203, 88–96.

Holcomb, P. J., Grainger, J., & Rourke, T. O. (2002). An Electrophysiological Study of the Effects of Orthographic Neighborhood Size on Printed Word Perception. *Journal of Cognitive Neuroscience*, 14(6), 938–950.

Holcomb, P.J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology*, 30, 47– 61.

Horner, V., Whiten, A., Flynn, E., & de Waal, F. B. M. (2006). Faithful replication of foraging techniques along cultural transmission chains by chimpanzees and children. *Proceedings of the National Academy of Sciences*, 103(37), 13878–13883.

Hruschka, D. J., Christiansen, M. H., Blythe, R. A., Croft, W., Heggarty, P., Mufwene, S. S., & Poplack, S. (2009). Building social cognitive models of language change. *Trends in Cognitive Sciences*, 13(11), 464–469.

Hudson Kam, C. I., & Newport, E. L. (2009). Getting it right by getting it wrong: when learners change languages. *Cognitive Psychology*, 59, 30-66.

Hudson Kam, C. L. H., & Newport, E. (2005). Regularizing Unpredictable Variation: The Roles of Adult and Child Learners in Language Formation and Change. *Language Learning and Development*, 1(2), 151–195.

Hurford, J. R. (2007). *The Origins of Meaning: Language in the Light of Evolution*. Oxford: Oxford University Press.

Indefrey, P. (2006). A meta-analysis of hemodynamic studies on first and second language processing: Which suggested differences can we trust and what do they mean? *Language Learning*, 56(SUPPL. 1), 279–304.

Izumi, S. (2003). Processing difficulty in comprehension and production of relative clauses by learners of English as a second language. *Language Learning*, 53(2), 285-323.

Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Jager, G. (2007). Evolutionary game theory and typology: a case study. *Language*, 83, 74-109.

Kaan, E. (2007). Event-Related Potentials and Language Processing: A Brief Overview. *Language and Linguistic Compass*, 1, 571–591.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. J. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, 15, 159–201.

Kaczer, L., Timmer, K., Bavassi, L., & Schiller, N. O. (2015). Distinct morphological processing of recently learned compound words: An ERP study. *Brain Research*, 1629, 309-317.

Kalish, M. L., Griffiths, T. L. & Lewandowsky, S. (2007). Iterated learning: intergenerational knowledge transmission reveals inductive biases. *Psychonomic Bulletin & Review*, 14, 288–294.

Kashima, Y. (2000). Maintaining cultural stereotypes in the serial reproduction of narratives. *Personality and Social Psychology Bulletin*, 26, 594–604.

Kaspar, F., & Schuster, H.G. (1987). Easily calculable measure for the complexity of spatiotemporal patterns. *Physical Review A*, 36, 842–848.

Kiefer, M., & Spitzer, M. (2000). Time course of conscious and unconscious semantic brain activation. *NeuroReport*, 11(11), 2401–2407.

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, 52(2), 205–225.

King, J. W., & Kutas, M. (1995). Who did what when? Using word-and clause- related ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, 7, 378–97.

Kirby, S. (2000). Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners. T*he Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*, 1–19.

Kirby, S. (2001). Spontaneous evolution of linguistic structure - An iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation*, 5(2), 102–110.

Kirby, S., & Hurford, J. R. (2002). The emergence of linguistic structure: An overview of the iterated learning model. In A. Cangelosi & D. Parisi (Eds.) *Simulating the Evolution of Language* (pp. 121-148). London, UK: Springer.

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

Kirby, S., Dowman, M., & Griffiths, T. L. (2007). Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences*, 104(12), 5241–5245.

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

Kirby, S., Smith, K., & Brighton, H. (2004). From UG to universals: Linguistic adaptation through iterated learning. *Studies in Language*, 28, 587–607.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141, 87–102.

Kolk, H. H. J., Chwilla, D. J., van Herten, M., & Oor, P. J. W. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language*, 85(1), 1–36.

Kroll, J.F., & Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33, 149–174.

Kuhl, P. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience*, 5, 831-843.

Kuperberg, G. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, 1146, 23-49.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, 12(12), 463–470.

Kutas, M., & Hillyard, A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, 207(4427), 203-205.

Kutas, M., & Hillyard, A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition*, 11, 539–50.

Lempel, A., & Ziv, J. (1976). On the complexity of finite sequences. *IEEE Transactions on Information Theory*, 22, 75-88.

Levinson, S. C., & Evans, N. (2010). Time for a sea-change in linguistics: Response to comments on 'The myth of language universals'. *Lingua*, 120, 2733-2758.

Levinson, S. C., & Gray, R. D. (2012). Tools from evolutionary biology shed new light on the diversification of languages. *Trends in Cognitive Sciences*, 16(3), 167–173.

Lewis, D. (1969). *Convention*. Cambridge, MA: Harvard University Press.

Lightfoot, D. (1997). Catastrophic change and learning theory. *Lingua*, 100, 171-192.

Lumaca, M., & Baggio, G. (2015). Cultural transmission and evolution of musical systems through iterated signaling games. *Submitted*.

Marchetto E., & Bonatti, L. L. (2013). Word and possible words in early language acquisition. *Cognitive Psychology*, 67, 130-150.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG and MEG data. *Journal of Neuroscience Methods*, 164, 177–190.

Martin, D., Hutchison, J., Slessor, G., Urquhart, J., Cunningham, S. J., & Smith, K. (2014). The Spontaneous Formation of Stereotypes via Cumulative Cultural Evolution. *Psychological Science*, 25(9), 1777-86.

McAvoy, A., & Hauert, C. (2015). Asymmetric Evolutionary Games. *PLOS Computational Biology*, 11(8).

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second- language word learning: minimal instruction produces rapid change. *Nature Neuroscience*, 7(7), 703–704.

Mesoudi, A., & Whiten, A. (2008). Review. The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 363(1509), 3489–3501.

Mestres-Misse, A., Rodriguez-Fornells, A., & Munte, T. F. (2007). Watching the brain during meaning acquisition. *Cerebral Cortex*, 8, 1858–1866.

Miller, G., A., Newman, E. B., & Friedman, E. A. (1958). Length-frequency statistics for written English. *Information and Control*, 1(4), 370-389.

Molinaro, N., Barber, H., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex*, 30, 1–23.

Monaghan, P. (2014). Age of acquisition predicts rate of lexical evolution. *Cognition*, 133(3), 530–534.

Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014). How arbitrary is language? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 369(1651).

Moreno, M., & Baggio, G. (2014). Role Asymmetry and Code Transmission in Signaling Games: An Experimental and Computational Investigation. *Cognitive Science*, 39(5), 918-43.

Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural Packaging in the Input to

Language Learning: Contributions of Prosodic and Morphological Marking of Phrases to the Acquisition of Language. *Cognitive Psychology*, 19, 498-550.

Morgan, J. L., & Newport, E. (1981). The Role of Constituent Structure in the Induction of an Artificial Language. *Journal of Verbal Learning and Verbal Behavior*, 20, 67-85.

Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience*, 24(4), 933-947.

Mueller, J. L., Hahne, A., Fujii, Y., & Friederici, A. D. (2005). Native and nonnative speakers' processing of a miniature version of Japanese as revealed by ERPs. *Journal of Cognitive Neuroscience*, 17, 1229–1244.

Mnte, T. F., Schiltz K., & Kutas, M.(1998). When temporal terms belie conceptual order. *Nature*, 395, 71–3.

Mnte, T., Matzke, M., & Johannes, S. (1997). Brain activity associated with syntactic incongruencies in words and pseudowords. *Journal of Cognitive Neuroscience*, 9(3), 318–329.

Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Bchel, C., & Weiller, C. (2003). Broca's area and the language instinct. *Nature Neuroscience*, 6(7), 774–781.

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, 20(2), 233–253.

Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically Based Sentence Processing Classes: Evidence from Event-Related Brain Potentials. *Journal of Cognitive Neuroscience*, 3(2), 151–165.

Newman-Norlund, R. D., Frey, S. H., Petitto, L.-A., & Grafton, S. T. (2006). Anatomical substrates of visual and auditory miniature second-language learning. *Journal of Cognitive Neuroscience*, 18, 1984–1997.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, 14, 11-28.

Nielsen, M., Cucchiaro, J., & Mohamedally, J. (2012). When the transmission of culture is child's play. *PLoS ONE*, 7(3).

Nieuwland, M. & van Berkum, M. (2005). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, 18, 1098–1111.

Niyogi, P. (2006). *The Computational Nature of Language Learning and Evolution*. Cambridge, MA: MIT Press.

Norris, J.M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta- analysis. *Language Learning*, 50, 417–528.

Nowak, M., A., & Komarova, N., L. (2001). Towards an evolutionary theory of language. *Trends in Cognitive Sciences*, 5(7), 288-295.

Nowak, M., A., Komarova, N. L., & Niyogi, P. (2001). Evolution of universal grammar. *Science*, 291(5501), 114–118.

Nowak, M., A., Komarova, N., L., & Niyogi, P. (2002). Computational and evolutionary aspects of language. *Nature*, 404, 495-498.

Nowak, M., A. (2006). Five rules for the evolution of cooperation. *Science*, 2134(5805), 1560.

Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after childhood: effects of proficiency. *Journal of Cognitive Neuroscience*, 17, 1212-28.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*.

Opitz B., & Friederici, A.D. (2004). Brain correlates of language learning: the neuronal dissociation of rule-based versus similarity based learning. *Journal of Neuroscience*, 24, 8436 – 8440.

Opitz, B., & Friederici, A. D. (2003). Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. *NeuroImage*, 19(4), 1730–1737.

Osterhout, L. (1997). On the brain response to syntactic anomalies: manipulations of word position and word class reveal individual differences. *Brain and Language*, 59(3), 494–522.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31(6), 785–806.

Osterhout, L., & Holcomb, P. J. (1993). Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes*, 8(4), 413–437.

Osterhout, L., McKinnon, R., Bersick, M., & Corey, V. (1996). On the Language Specificity of the Brain Response to Syntactic Anomalies: Is the Syntactic Positive Shift a Member of the P300 Family? *Journal of Cognitive Neuroscience*, 8(6), 507–526.

Osterhout, L., McLaughlin, J., Kim, A., Greenwald, R., & Inoue, K. (2004). Sentences in the brain: Event-related potentials as real-time reflections of sentence comprehension and language learning. In M. Carreiras & C. Clifton, Jr. (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP, and beyond*. Psychology Press.

Osterhout, L., McLaughlin, L., Pitknen, I., Frenck-Mestre C., & Molinaro, N. (2006). Novice learners, longitudinal designs, and event-related potentials: a means for exploring the neurocognition of second language processing. *Language Learning*, 56(1), 199–230.

Page, E. B. (1963). Ordered hypotheses for multiple treatments: A significance test for linear ranks. *Journal of the American Statistical Association*, 58(301), 216–230.

Pagel, M., Atkinson, Q. D., & Meade, A. (2007). Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature*, 449(7163), 717–720.

Pagel, M., Atkinson, Q. D., S Calude, A., & Meade, A. (2013). Ultraconserved words point to deep language ancestry across Eurasia. *Proceedings of the National Academy of Sciences of the United States of America*, 110(21), 8471–6.

Perani, D., Paulesu, E., Sebastian-Galles, N., Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S., Fazio, F., & Mehler, J. (1998). The bilingual brain: Proficiency and age of acquisition of the second language. *Brain*, 121, 1841-1852.

Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, 15(2), 202-206.

Perani, D., & Abutalebi, J. (2015). Bilingualism, dementia, cognitive and neural reserve. *Current Opinion in Neurology*, 28(6), 618-625.

Perfetti, C., Wlotko, E. W., & Hart, L. A. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology – Learning, Memory and Cognition*, 31 (6), 1281–1292.

Perfors, A., & Navarro, D. J. (2014). Language evolution can be shaped by the structure of the world. *Cognitive Science*, 38(4), 775–793.

Perfors, A., Tenenbaum, J.B, Regier, T. (2011). The learnability of abstract syntactic principles. *Cognition*, 118, 306-338.

Perruchet, P., & Rey, A. (2005). Does the mastery of center-embedded linguistic structures distinguish human from nonhuman primates? *Psychonomic Bulletin & Review*, 12(2), 307-313.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends in Cognitive Sciences*, 10(5), 232-238.

Piaget, J. (1926). *The Language and Thought of the Child*. New York: Harcourt Brace & Company.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *The Behavioral and Brain Sciences*, 27(2), 169–190; discussion 190–226.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Pinker, S., & P. Bloom. (1990). Natural language and natural selection. *The Behavioral and Brain Sciences*, 13, 707–784.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118, 2128–2148.

Pothos, E. M. (2007). Theories of Artificial Grammar Learning. *Psychological Bulletin*, 133(2): 227-244.

Puglisi, A., Baronchelli, A., & Loreto, V. (2008). Cultural route to the emergence of linguistic categories. *Proceedings of the National Academy of Sciences of the United States of America*, 105(23), 7936–7940.

Ramnani, N., & Miall, R. C. (2004). A system in the human brain for predicting the actions of others. *Nature Neuroscience*, 7(1), 85–90.

Ramscar, M., Gitcho, N. (2007). Developmental change and the nature of learning in childhood. *Trends in Cognitive Sciences*, 11, 274-279.

Reali, F., & Christiansen, M. H. (2007). Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language*, 57, 1-23.

Reali, F., & Griffiths, T. L. (2009). The evolution of frequency distributions: Relating regularization to inductive biases through iterated learning. *Cognition*, 111(3), 317–328.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology*, 81(1), 115-119.

Regel, S., Meyer, L., & Gunter, T. C. (2014). Distinguishing Neurocognitive Processes Reflected by P600 Effects: Evidence from ERPs and Neural Oscillations. *PLoS ONE*, 9(5).

Rendell, L., Fogarty, L., Hoppitt, W. J. E., Morgan, T. J. H., Webster, M. M., & Laland, K. N. (2011). Cognitive culture: Theoretical and empirical insights into social learning strategies. *Trends in Cognitive Sciences*, 15(2), 68–76.

Richerson, P. & Boyd, R. (2005). *Not by Genes Alone: How Culture Transformed Human Evolution*. Chicago: University of Chicago Press.

Ringe, D., & Eska, J.F. (2013). *Historical Linguistics*. Cambridge: Cambridge University Press.

Rohrmeier, M., Fu, Q., & Dienes, Z. (2012). Implicit learning of recursive context-free grammars. *PLoS ONE,* 7(10).

Rsler, F., Ptz, P., Friederici, A., & Hahne, A. (1993). Event-Related Brain Potentials While Encountering Semantic and Syntactic Constraint Violations. *Journal of Cognitive Neuroscience*, 5(3), 345–362.

Saffran, J. R., Newport, E. L. & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35, 606-621.

Saloni, Z. (2007). *Czasownik polski: Odmiana, slownik*. Warszawa: Wydawnictwo Naukowe PWN.

Schacht, A., Sommer, W., Shmuilovich, O., Martenz, P. C., & Martn-Loeches, M. (2014). Differential Task Effects on N400 and P600 Elicited by Semantic and Syntactic Violations. *PLoS ONE*, 9(3).

Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, 21, 211-232.

Scott-Phillips, T. C., & Kirby, S. (2010). Language evolution in the laboratory. *Trends in Cognitive Sciences*, 14(9), 411–417.

Scott-Phillips, T. C., Kirby, S., & Ritchie, G. R. S. (2009). Signalling signalhood and the emergence of communication. *Cognition*, 113(2), 226–233.

Seidenberg, M. S., & Gonnerman, L. M. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences*, 4(9), 353-361.

Selten, R., & Warglien, M. (2007). The emergence of simple languages in an experimental coordination game. *Proceedings of the National Academy of Sciences of the United States of America*, 104(18), 7361–7366.

Silvey, C., Kirby, S., & Smith, K. (2014). Word Meanings Evolve to Selectively Preserve Distinctions on Salient Dimensions. *Cognitive Science*, 39(1), 212–226.

Singleton, J. L., & Newport, E. L. (2004). When learners surpass their models: the acquisition of American sign language from inconsistent input. *Cognitive Psychology*, 49, 370-407.

Skyrms, B. (2009). Evolution of signalling systems with multiple senders and receivers. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364(1518), 771–779.

Skyrms, B. (2010). *Signals. Evolution, learning and information.* Oxford, UK: Oxford University Press.

Slobin, D. I. (1997). *The cross-linguistic study of language acquisition.* Hillsdale, NJ: Lawrence Erlbaum.

Slobin, D. I. (2005). From ontogenesis to phylogenesis: What can child language tell us about language evolution. In Langer, J., Parker, S. T., & Milbrath, C. (Eds.), *Biology and knowledge revisited: From neurogenesis to psychogenesis* (pp.255-285). Mahwah, NJ: Lawrence Erlbaum Associates.

Slobin, D. I., & Bever, T. G. (1982). Children use canonical sentence schemas: A cross-linguistic study of word order and inflections. *Cognition*, 12, 229-265.

Smith, K. (2004). The evolution of vocabulary. *Journal of Theoretical Biology*, 228(1), 127–142.

Smith, K., & Wonnacott, E. (2010). Eliminating unpredictable variation through iterated learning. *Cognition*, 116(3), 444–449.

Smith, K., & Kirby, S. (2008). Cultural evolution: implications for understanding the human language faculty and its evolution. *Philosophical Transactions of the Royal Society B*, 363(1509), 3591-3603.

Smith, K., Kirby, S., & Brighton, H. (2003). Iterated learning: a framework for the emergence of language. *Artificial Life*, 9(4), 371–386.

Smith, K., Tamariz, M., & Kirby, S. (2013). Linguistic structure is an evolutionary trade-off between simplicity and expressivity. *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*, 1348–1353.

Snijders, T.M., Vosse, T., Kempen, G., Van Berkum, J. J., Petersson, K.M., & Hagoort, P. (2009). Retrieval and unification of syntactic structure in sentence comprehension: an fMRI study using word-category ambiguity. *Cerebral Cortex*, 19, 1493-1503.

Soliz, J., & Giles, H. (2014). Relational and identity processes in communication: a contextual and meta-analytical review of communication accommodation theory. In: E. Cohen (Ed.), *Communication Yearbook 38* (pp. 106–143). Thousand Oaks, CA: Sage.

St. Clair, M. C., Monaghan, P., & Ramscar, M. (2009). Relationships between language structure and language learning: The suffixing preference and grammatical categorization. *Cognitive Science*, 33, 1317-1329.

Steels, L. (2003). Evolving grounded communication for robots. *Trends in Cognitive Sciences*, 7, 308–312.

Stein, M., Dierks, T., Brandeis, D., Wirth, M., Strik, W., & Koenig, T. (2006). Plasticity in the adult language system: A longitudinal electrophysiological study on second language learning. *Neuroimage*, 33(2), 774–783.

Steinhauer, K., & Connolly, J.F. (2008). Event-related potentials in the study of language. In: B. Stemmer & H. A. Whitaker (Eds.), *Handbook of the neuroscience of language* (pp. 91-103). Amsterdam: Elsevier.

Steinhauer, K., & Drury, J. A., (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain Language*, 120, 135–162.

Steinhauer, K., White, E.J., & Drury, J.E. (2009). Temporal dynamics of late second language acquisition: evidence from event-related brain potentials. *Second Language Research*, 25 (1), 13–41.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences of the United States of America*, 106(26), 10587–10592.

Stolk, A., Noordzij, M. L., Volman, I., Verhagen, L., Overeem, S., Van Elswijk, G., & Toni, I. (2014). Understanding communicative actions: A repetitive TMS study. *Cortex*, 51(1), 25–34.

Swarup, S., & Gasser, L. (2009). The Iterated Classification Game: A New Model of the Cultural Transmission of Language. *Adaptive Behavior*, 17(3), 213–235.

Tallerman, M. (2007). Did our ancestors speak a holistic protolanguage? *Lingua*, 117(3), 579–604.

Tamariz, M., & Smith, A. D. M. (2008). Regularity in mappings between signals and meanings. *The Evolution of Language: Proceedings of the 7th International Conference*, 315–322.

Tamariz, M., Brown, J. E., & Murray, K. (2010). The role of practice and literacy in the evolution of linguistic structure. *The Evolution of Language: Proceedings of the 8th International Conference*, 313–320.

Tan, R., & Fay, N. (2011). Cultural transmission in the laboratory: Agent interaction improves the intergenerational transfer of information. *Evolution and Human Behavior*, 32(6), 399–406.

Tettamanti, M., Alkadhi, H., Moro, A., Perani, D., Kollias, S., & Weniger, D. (2002). Neural correlates for the acquisition of natural language syntax. *Neuroimage*, 17(2), 700-9.

Theisen-White, C., Kirby, S., & Oberlander, J. (2011). Integrating the horizontal and vertical cultural transmission of novel communication systems. Expanding the Space of Cognitive Science. *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, 956–961.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition.* Boston: Harvard University Press.

Tomasello, M. (2008). *Origins of human communication.* Cambridge, MA: MIT Press.

Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition*, 92, 231–270.

Valian, V., & Coulson, S. (1988). Anchor points in language learning: the role or marker frequency. *Journal of Memory and Language*, 27, 71-86.

Valian, V., & Levitt, A. (1996). Prosody and Adults' Learning of Syntactic Structure. *Journal of Memory and Language*, 35, 497-516.

Van Herten, M., Kolk, H. H. J., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*, 22(2), 241–255.

Van Petten, C. K., & Kutas, M. (1991). Influences of semantic and syntactic context in open-and closed-class words. *Memory and Cognition*, 19, 95–112.

Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning Memory & Cognition*, 25, 394–417.

Vaughan-Evans, A., Kuipers, J. R., Thierry, G., & Jones, M. W. (2014). Anomalous Transfer of Syntax between Languages. *Journal of Neuroscience*, 34(24), 8333–5.

Vogt, P. (2005). The emergence of compositional structures in perceptually grounded language games. *Artificial Intelligence*, 167(1-2), 206–242.

Weber-Fox, C. M., & Neville, H. J. (1996). Maturational Constraints on Functional Specializations for Language Processing: ERP and Behavioral Evidence in Bilingual Speakers. *Journal of Cognitive Neuroscience*, 8(3), 231–256.

Whiten, A., & Flynn, E. (2010). The transmission and evolution of experimental microcultures in groups of young children. *Developmental Psychology*, 46(6), 1694–1709.

Whiten, A., Spiteri, A., Horner, V., Bonnie, K. E., Lambeth, S. P., Schapiro, S. J., & de Waal, F. B. M. (2007). Transmission of Multiple Traditions within and between Chimpanzee Groups. *Current Biology*, 17(12), 1038–1043.

Willems, R.M., de Boer, M., de Ruiter, J.P., Noordzij, M.L., Hagoort, P., & Toni, I. (2010). A dissociation between linguistic and communicative abilities in the human brain. *Psychological Science*, 21, 8-14.

Wilson, C. (2006). An experimental and computational study of velar palatalization. *Cognitive Science*, 30, 945–982.

Wilson, D., & Sperber, D. (2002). Truthfulness and relevance. *Mind*, 111, 583-632.

Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., Czigler, I., Csepe, V., Ilmoniemi, R.J. & Ntnen, R. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology*, 36, 638–42.

Wray, A. (1998). Protolanguage as a holistic system for social interaction. *Language and Communication*, 18, 47–67.

Yang, C. (2004). Universal Grammar, Statistics, or Both. *Trends in Cognitive Sciences*, 8(10), 451-456.