Processing and representation of linear and hierarchical syntactic dependencies in an artificial language learning paradigm

ROTONDI, Irène

Abstract

In this thesis we used an artificial language learning paradigm to investigate the learning and generalization of different types of rules based on linear and hierarchical syntactic dependencies. We also explored which factors (i.e., type of units, phonological, prosodic and semantic cues) influenced the learning. In all experiments participants were presented with grammatical sentences and their ability to extract syntactic regularities was studied during test phases. In the first two experiments, the rule was based on the linear position of the elements in the sentence. In the other experiments (from experiment 3 to 6), the rule was based on the position of the elements in the hierarchical structure of the sentence. While experiments 1, 2, and 3 did not revealed learning, the subsequent experiments showed that the presence of specific phonological and prosodic cues (e.g., syllables number, pauses) facilitates rule learning. Interestingly, also semantic cues seem to play a role when high level processing was required to generalize the knowledge of the rule to novel syntactic structures.

Reference


URN : urn:nbn:ch:unige-189490
DOI : 10.13097/archive-ouverte/unige:18949
Sous la co-direction du Professeur Ulrich H. Frauenfelder et du Docteur Julie Franck

PROCESSING AND REPRESENTATION OF LINEAR AND HIERARCHICAL SYNTACTIC DEPENDENCIES IN AN ARTIFICIAL LANGUAGE LEARNING PARADIGM

THESE

Présentée à la Faculté de psychologie et des sciences de l'éducation de l'Université de Genève pour obtenir le grade de Docteur en Psychologie

par

Irene Maria Cecilia ROTONDI

de

Italie

Thèse No 501

GENEVE

Février, 2012

No étudiant: 06-329-270
ACKNOWLEDGMENTS

This dissertation would not have been possible without the help and support of many important people. I am grateful to Uli and Julie, who gave me the time and resources to deepen my knowledge and experience in research.

Uli, I would like to thank you for your guidance all along the challenging way of this PhD. Thank you for your great patience, especially during the preparation of this manuscript, and for always reminding me the importance of being synthetic and clear in the presentation of my ideas. Julie, thank you for your detailed and constructive comments, for your guidance and advice throughout this project.

I would like to express my gratitude to Doug Saddy for the interest demonstrated in this project from its very first steps and for inspirational discussions regarding these experiments. My gratitude goes also to the other members of my jury, Pascal Zesiger and Luigi Rizzi, for accepting to read and evaluate this thesis.

I am indebted to Andres Posada for the good hints he gave me during the preparation of the experiments with the software Presentation, and to Jean-Philippe Goldman for his help and the precious scripts that facilitated my work with Praat and MBROLA. I am also grateful to Simone Giani for his help in the creation of the music stimuli for the first experiment.

I had the chance to have wonderful colleagues and friends who made my time in Geneva enjoyable and productive. First, I would like to thank Jennifer Martin and Audrey Bürki. Jennifer, thank you for your kindness and your optimism. You have always been available to answer my many questions and support me with chocolate and good advice. I will always be in debt with you of too many “verres”. Audrey, I have always admired your determination and energy. Your dedication and dynamism showed me the person I want to be when I grow up. I could not have wished for better colleagues.

Many thanks also to Natalia Kartushina, Hélène Delage, Romy Lassotta, and Odile Bagou for their support, for the coffee breaks that allowed us to discuss our projects and share useful
information. I am also grateful to Akira Omaki for reading and discussing with me parts of this manuscript.

I am grateful to Marina Laganaro and the NeuroPsychoLinguistics group who allowed me to collaborate to their research projects. Even if it was only for a short period of time, I appreciated your welcome and to work with you. I would also like to thank the members of the Processing, Acquisition and Disorders of language group (a.k.a. Les filles de la logo) for sharing nice lunches and interesting seminars.

I would like to thank my “Geneva family”, Martijn, Marcello, Chiara, Cristina, Nele, Joanna and the “CISA people”, you all made my years in Geneva happy and fun. Martijn, thank you for your friendship and for the helpful discussions about my project. Cristina, I am grateful to have such a wise friend who helped me coping with the challenges of the PhD student life. Thank you for revising this manuscript, you are also very brave. Chiara, Nele and Marcello, if it is true that (he) who finds a friend finds a treasure, finding the three of you was a true fortune.

Lastly, I am immensely grateful to my family who supported me in all my pursuits, encouraged me and taught me the importance of accomplish a task although difficult. Giacomo, thank you for your patience and for coping with my mood swings. Thank you for your love and encouragement.
# TABLE OF CONTENTS

INTRODUCTION .............................................................................................................................................. 1

CHAPTER 1. THEORETICAL FRAMEWORK

1.1. INTRODUCTION ........................................................................................................................................... 7

1.2. LANGUAGE ACQUISITION .............................................................................................................................. 8
  1.2.1. THE PUZZLE OF LANGUAGE ACQUISITION ......................................................................................... 8
    1.2.1.1. Universal Grammar Theory ........................................................................................................ 10
    1.2.1.2. Usage-based Theory ..................................................................................................................... 15
  
1.3. SYNTACTIC PROCESSING IN NATURAL LANGUAGE ...................................................................................... 18
  1.3.1. THE ROLE OF SYNTACTIC STRUCTURE IN SYNTACTIC PROCESSING ............................................ 20
  1.3.2. THE ROLE OF SEMANTICS IN SYNTACTIC PROCESSING .................................................................. 22
  1.3.3. THE ROLE OF PROSODY IN SYNTACTIC PROCESSING .................................................................... 24

1.4. ARTIFICIAL LANGUAGE LEARNING PARADIGM ......................................................................................... 26
  1.4.1. METHODOLOGY ...................................................................................................................................... 28
    1.4.1.1. Artificial Languages ...................................................................................................................... 28
    1.4.1.2. Experimental tasks ......................................................................................................................... 31
    1.4.1.3. Population ...................................................................................................................................... 34
  1.4.2. REVIEW OF THE LITERATURE IN ALL ................................................................................................. 36
    1.4.2.1. ALL in the linguistic domain ......................................................................................................... 36
    1.4.2.2. ALL in the non-linguistic domains ................................................................................................. 45
    1.4.2.3. Learning Constraints ...................................................................................................................... 47
    1.4.2.4. Generalization ................................................................................................................................. 51
    1.4.2.5. Types of Knowledge ......................................................................................................................... 57
    1.4.2.6. Summary ......................................................................................................................................... 63

1.5. RESEARCH QUESTIONS AND EMPIRICAL FRAMEWORK ........................................................................... 64
  1.5.1. RESEARCH QUESTIONS .......................................................................................................................... 64
  1.5.2. EMPIRICAL FRAMEWORK ..................................................................................................................... 69
    1.5.2.1. Types of Rule ................................................................................................................................... 70
    1.5.2.2. Types of Knowledge ....................................................................................................................... 73
  1.5.3. CONSIDERATIONS ABOUT DATA ANALYSES ..................................................................................... 74
    1.5.3.1. D-prime ........................................................................................................................................... 75
    1.5.3.2. Template Analysis ........................................................................................................................... 77
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.6.1. RESULTS</td>
<td>168</td>
</tr>
<tr>
<td>3.6.2. DISCUSSION</td>
<td>169</td>
</tr>
</tbody>
</table>

**CHAPTER 4. THE ROLE OF SEMANTIC CUES**

4.1. INTRODUCTION ................................................................. 173

4.2. EXPERIMENT 6 ................................................................. 177
   4.2.1. METHOD .................................................................. 177
   4.2.2. RESULTS ............................................................... 179
   4.2.3. DISCUSSION ........................................................... 184

4.3. COMPARISONS BETWEEN EXPERIMENT 5 AND EXPERIMENT 6 ............... 185
   4.3.1. RESULTS ............................................................... 185
   4.3.2. DISCUSSION ........................................................... 186

**CHAPTER 5. GENERAL DISCUSSION**

5.1. SUMMARY OF THE DISSERTATION ............................................... 191

5.2. ARE ADULTS ABLE TO LEARN LINEAR DEPENDENCIES IN AN ARTIFICIAL LANGUAGE? ........................................................................ 193

5.3. ARE ADULTS ABLE TO LEARN HIERARCHICAL DEPENDENCIES IN AN ARTIFICIAL LANGUAGE? ............................................................ 195
   5.3.1. PHONOLOGICAL AND PROSODIC CUES ......................... 198
   5.3.2. SEMANTIC CUES ..................................................... 199
   5.3.3. SUMMARY .............................................................. 200

5.4. NATURE OF THE ACQUIRED KNOWLEDGE .................................. 201

5.5. METHODOLOGICAL ISSUES ...................................................... 204

5.6. FUTURE PERSPECTIVES ......................................................... 208

REFERENCES ................................................................................. 211

LIST OF APPENDIXES .................................................................... 231

FRENCH SUMMARY ........................................................................ 243
INTRODUCTION

Language acquisition, and in particular the acquisition of syntax, is at first glance one of the most difficult learning tasks which infants have to face. However, children learn to speak without any apparent effort and in a limited amount of time. It has been claimed (Chomsky, 1980) that the information available in the input is not sufficient for children to infer the correct grammar. How do children reach grammatical knowledge? Two main theoretical approaches, Universal Grammar theory and Usage-based theory, try to answer this question studying the mechanisms that learners exploit to acquire linguistics knowledge. The Universal Grammar theory assumes that infants are endowed with innate constraints that allow them to select among all possible grammars the correct one. According to the Usage-based theory, infants would not be innately equipped with abstract linguistic knowledge, instead they would move gradually from a concrete knowledge of word and expressions to more abstract schematic constructions proper of adult grammar knowledge. Evidence supporting one approach or the other came not only from results of experiments employing natural language stimuli but also from artificial language learning paradigms.

In the last 50 years, the artificial language learning paradigm has been extensively used to investigate the cognitive processes involved in language acquisition. Artificial languages mimic aspects of natural language structures and explore target features (e.g., segmentation, categories identification, etc.) in isolation. In the literature, some studies focused primarily on the mechanisms used by infants and adults to perform lexical segmentation; other studies focused on syntactic relationships learning. Our work can be inscribed on this second group of studies since it aims to investigate the learning of syntactic dependencies. With this work we also aim to investigate phonological, prosodic and semantic factors influencing grammatical rule learning.

In order to contribute to the study of syntactic dependencies learning we will employ two types of rules characterizing our artificial languages. The first rule will be characterized by linear syntactic dependencies based on the position of the elements in the linear sequence of the items. The second rule will be characterized by hierarchical syntactic dependencies based on the position of the elements in the hierarchical structure of the sentence. In all
experiments participants will be presented with training phases in which they will listen to sentences characterized by the syntactic dependency they have to learn. Subsequently, they will perform test phases in which their ability to extract syntactic regularities will be studied. Two main purposes characterize the present dissertation: the first one is to collect evidence about the learnability of different types of rules based on linear and hierarchical syntactic dependencies in an artificial language; the second purpose is to investigate which factors influence rule learning.

Chapter 1 will lay the theoretical foundations of this dissertation. We will outline the problem underlying language acquisition and present the two main theoretical approaches accounting for children language acquisition, namely the Universal Grammar Theory and the Usage-based Theory. Subsequently, we will provide a review of the models explaining syntactic parsing. In particular, we will focus on the role of syntactic structure, prosody and semantic in syntactic processing. Finally, we will review the literature concerning artificial language learning paradigms and present the Research Questions characterizing this dissertation.

In Chapter 2, we will present two experiments investigating the acquisition of linear syntactic dependencies. Experiment 1 will be conducted in the music domain while Experiment 2 will be conducted in the linguistic one. The aim of these two experiments is to explore adults’ capacity to learn linear dependencies and to examine differences and similarities between the learning of regularities involving different types of units. In Chapter 3, we will investigate hierarchical syntactic dependencies learning in the linguistic domain with focus on the influence of prosodic and phonological cues. Subsequently, in Chapter 4, the role of semantic cues in the learning of the same type of dependencies will be studied. The aim of these four experiments is to examine whether human adults are able to learn syntactic dependencies in an artificial language when they are characterized as hierarchical.

The last chapter of this dissertation will include a synthesis of the main results of the experiments and their relevance in the context of language acquisition and artificial language learning. The nature of the knowledge resulting from the learning task will also be discussed.
Finally, we will highlight few methodological issues of our work and propose directions for future research in artificial language learning.
Chapter 1. Theoretical Framework
1.1. Introduction

In the current chapter we will overview the theoretical framework in which our empirical work is situated. This introduction will allow us to move from more general topics, such as the theories accounting for language acquisition, to more specific topics, such as the types of information influencing syntactic processing. Subsequently, we will outline the contribution of artificial language learning paradigms to the study of the mechanisms that enable natural language learning. Finally, in the last part of the chapter, we will outline the research goals of this dissertation which are (a) investigating adults’ capacity to learn agreement rules in the context of an artificial language and (b) investigating the factors influencing rule acquisition in an artificial language learning paradigm. We will also present the empirical framework of our work which comprises the aspects common to all the experiments, namely the types of rule implemented in the artificial language, the types of knowledge tested, and some considerations about data analysis.

When infants acquire language they are faced with a difficult task. In fact, as we will present in Section 1.2.1., the characteristics of the input they are exposed to cannot account for the linguistic knowledge they acquire. How do children reach linguistic knowledge? Two main theoretical approaches try to answer this question studying the mechanisms that learners exploit to acquire linguistics knowledge. In the framework of the Universal Grammar theory (Section 1.2.1.1.), learners are supposed to be pre-equipped with a system of principles (that represent the invariant properties of the languages) and their task consists in setting the value of parameters (that are binary choices that determine the variability among languages) according to the input they are exposed to. On the other hand, in the framework of the Usage-based theory (Section 1.2.1.2.), children would not be endowed with any innate abstract linguistic knowledge. Instead, they would learn concrete linguistic expressions by imitation and then move on to more abstract representations of their linguistic knowledge.

Once learners have acquired a language, sentence comprehension seems to be an easy task. However, its characterization is still a matter of debate. In fact, it is still unclear at what stage (i.e., when in the process) the information of different types (e.g., syntactic, semantic, prosodic) may influence sentence comprehension. Section 1.3. will be devoted to the review of
the main approaches that have been proposed to characterize sentence processing: modular accounts and interactive accounts. The former assume that in the first stages of the processing only syntactic information plays a role, while the latter acknowledge the influence of other types of information from early on.

While the research concerning syntactic processing mainly used experimental paradigms involving sentence disambiguation, the investigation of the mechanisms playing a role in language acquisition has extensively exploited artificial language learning paradigms (Section 1.4.). The principle behind this kind of studies is to use simplified systems to investigate different aspects of processing independently of one another. Artificial language learning experiments have been employed to study different topics (e.g., word segmentation, word order, grammatical relations, syntactic categories) and in different experimental settings. The review of the methodology and of the main results in the literature about artificial language learning will allow us to have a clear picture of the state of the art and to introduce the research questions characterizing this dissertation.

1.2. Language Acquisition

In the present section we will outline the theoretical framework for the studies in language acquisition. The domain is indeed very broad; therefore, we will first consider the arguments concerning the question of how children reach linguistic knowledge from the input they are exposed to. Subsequently, we will introduce two main theoretical approaches that try to account for the mechanisms that learners exploit in order to acquire linguistic knowledge.

1.2.1. The Puzzle of Language Acquisition

Children learn to speak without any apparent effort and in a limited amount of time even if language acquisition is a difficult learning task. At 6 months they start babbling, that is, they produce sequences of repetitive syllables; at 10 – 12 months infants produce their first words and at 20 – 24 months they put words together in simple sentences. By the age of 4 infants are able to master almost all syntactic constructions and possess a rich lexicon. An ongoing debate attempts to define what kind of information infants are able to extract from the
input, how this is done, and whether they are equipped with innate constraints. We introduce this issue by means of an example. We can consider how infants infer the rule that guides the transformation of a declarative sentence to an interrogative one. Children’s input is mostly characterized by sentences like the following:

\[(1a) \quad \text{Bill can sing} \quad \quad (1b) \quad \text{Can Bill } _{\_} \text{ sing?} \]
\[(2a) \quad \text{Bill is nice} \quad \quad (2b) \quad \text{Is Bill } _{\_} \text{ nice?} \]

One rule that could be inferred from the regularities in this input is: “to produce an interrogative sentence move the verb to the first position”. Consider now sentence (3a):

\[(3a) \quad \text{Bill, who is nice, can sing} \quad \quad (3b) \quad \ast \text{Is Bill who } _{\_} \text{ nice can sing?} \]

If such rule was inferred, we would then expect children to produce structures like (3b). However, this is not the case (Chomsky, 1968; Pinker, 1994). The knowledge concerning the relation between declarative sentences and the corresponding interrogative sentences is one of the arguments of the Poverty of the Stimulus.

The Poverty of the Stimulus (POS) argument and its formalization has been at first attributed to Chomsky (1980). However, as Pullum and Scholz pointed out in their paper (Pullum & Scholz, 2002), the POS argument has not been defined univocally in the literature. Rather than a single argument, it could be considered a collection of “mutually supporting claims” (Laurence & Margolis, 2001, p. 222). In the attempt to simplify the general assumption of these arguments we refer to the characterization proposed by Pullum and Scholtz (2002, p. 18):

a) Human infants learn their first languages either by data-driven learning or by innately-primed learning.
b) If human infants acquire their first languages via data-driven learning, then they can never learn anything for which they lack crucial evidence.
c) But infants do in fact learn things for which they lack crucial evidence.
d) Thus human infants do not learn their first languages by means of data-driven learning.
e) Conclusion: human infants learn their first languages by means of innately-primed learning.
Two facts determine the characteristics of the evidence infants are exposed to: 1) the input contains only a finite set of the infinite possible sentences in natural language; 2) the input does not contain sentences that falsify the incorrect hypothesis (Cowie, 2010). Since the input is not sufficient for the infant to learn the grammar, language cannot be learned from the input. However, despite the characteristics of the input, infants succeed in language acquisition converging on the grammar of their linguistic community. Interestingly, across languages infants show universal similarities in the acquisition process¹.

How do children reach grammatical knowledge? Two main theoretical approaches try to answer this question studying the mechanisms that learners exploit to acquire linguistic knowledge. In the following sections we will introduce the main aspects of these positions.

1.2.1.1. Universal Grammar Theory

It has been claimed (see previous section) that the information available in the input is not sufficient for children to infer the correct grammar of their language. For this reason, Chomsky (1959, 2000) proposed that infants are equipped with constraints that allow them to select among all possible grammars the correct one. This theory, called Universal Grammar, developed from the observation that human languages, despite the differences in structural properties, share characteristics at an abstract level. Initially, the idea of a Universal Grammar was considered as a theory of the rules and the conditions in which those rules were applied. The child has an initial set of possible grammars; her task is to discover the rules specific to the language she is exposed to. The introduction of the concept of parameter (Chomsky, 1965) allowed the evolution towards a different concept of language acquisition, namely acquisition as parameter setting. In this context, Universal Grammar is conceived as a system of principles (e.g., structure-dependency, binding, etc.) and parameters (e.g., the head direction parameter, the pro-drop parameter, etc.). Principles represent the invariant properties of languages, while parameters are binary choices which underlie the variability among languages. As Rizzi (2007) pointed out, the introduction of the notion of parameter changed the view of language acquisition. Learners are supposed to be pre-equipped with principles

¹ For instance, both in spoken and sign language they begin babbling at 6-8 months (Petitto & Marentette, 1991), and at around the age of 2 they use infinite verbs in main clauses. This phenomenon has been observed in Danish, Dutch, French, German, Russian and Swedish (e.g., Haegeman, 1995; Hamann & Plunkett, 1998; Wexler, 1994).
and parameters, and their task consists in setting the value of the latter according to the input they are exposed to. Learning is considered as the selection of the appropriate values of the parameters among the possibilities. This selection is based on linguistic experience. Therefore, the grammar of a specific language is the Universal Grammar with the parameters fixed by certain values. As a result, in order to study syntax acquisition it is necessary to explore the progressive fixation of the parameters by the infant.

All languages are characterized by structural flexibility, which allows them to organize a finite set of elements into an infinite set of combinations. This is possible since words are grouped together to form phrases, which in turn form larger phases, and so on. These groups of words are called constituents and represent the basic units of the sentence. It is possible to identify these units by means of “constituency tests”. In other words, these tests allow us to pick up a part of a sentence and verify whether it is a phrase or not. For instance, in the test called “cleft formation test”, the chosen part has to be used to create a novel sentence of the type “It is [chosen part] that...” followed by the original sentence. Let us illustrate this test using an example. The original sentence is The pirates hid many treasures in the ship. We choose the part many treasures. Using the cleft formation test we obtain the sentence It is [many treasures] that the pirates hid in the ship, which is a grammatical sentence in English. Therefore, we can conclude that many treasures is a constituent of the sentence. We choose now another part of the sentence: treasures in. The sentence we obtain is *It is [treasures in] that the pirates hid many the ship. This sentence is not acceptable, therefore treasures in is not a constituent. There are several other tests implying other procedures such as substitution, coordination or deletion. All these tests share the same aim of finding the basic syntactic units of the sentence.

Empirical evidence showed that infants are indeed able to identify and represent constituents. This is possible since there are prosodic cues that facilitate the identification of constituents’ boundaries. For instance, as summarized by Shattuck-Hufnagel and Turk (1996), some acoustic phenomena such as intonation, phoneme lengthening and pauses tend to occur at syntactic boundaries showing correspondence between syntax and prosody. Gerken and collaborators (Gerken, Jusczyk, & Mandel, 1994) performed an experiment in which 9-month old infants were exposed to natural speech samples in which artificial pauses were added. In
one condition the pause was added after the subject and in the other condition it was added after the verb. Infants were tested in a preferential listening paradigm. Results showed that infants listened longer to samples containing the pause after the subject. The authors interpreted these results as evidence that infants detected other prosodic cues such as vowel lengthening and intonation marking syntactic boundaries. In contrast, infants rejected the sentences in which the pause was inserted after the verb, as it did not coincide with any other cue.

So far we have reviewed some evidence about the identification of constituents. In the remainder of this section we will focus on the way constituents are organized in sentences.

Constituents are organized following hierarchical organization (Chomsky, 1957), which is a principle of the Universal Grammar. Hierarchical organization allows the combination of basic units into more complex patterns. Lower-level elements (e.g., constituents) can be combined to form higher-level structures (e.g., clauses). This combination can generate infinitely long and complex grammatical sentences, as illustrated in (4):

(4) The boy [saw the girl].
    My friends [said [that [the boy [saw the girl]]]].
    My friends [said [that [the boy [saw [the girl [who was wearing [the pink dress]]]]]]].

Sentence structural relations are not based on the linear order of the words (e.g., the subject always immediately precedes the verb), but rather on the hierarchical organization of the constituents. In other words, the distribution of the words in the sentence is not determined by rules mapping the position of an element as a function of the preceding or following one, but rather by rules mapping the position of the elements in the hierarchical structure of the sentence.

Empirical evidence showed that children interpret sentences relying on their hierarchical structure. A relevant study is the one performed by Lidz and Musolino (2002). Using ambiguous sentences containing quantifiers and negation (e.g., The detective didn’t find two guys), they investigated whether children have linguistic representations of the hierarchical structure and of the relations over this structure. Crucially, they compared the performance of English-speaking children and that of Kannada-speaking children (a Dravidian
language spoken in the south-western India). This comparison allowed them to tease apart linear order from hierarchical structure. In English, negation both precedes and c-commands the object, while in Kannada, negation c-commands the object, but does not precede it. Results have shown that Kannada children interpret the sentences in the same way as English children; this observation indicates that the interpretation is conditioned by the hierarchical position of the negation, rather than by its linear position.

Let us now consider an example of a parameter: the Head direction one. This parameter characterizes the word order in different languages. For instance, English is a head-initial language, which means that the head precedes the complement in the phrase, whereas Japanese is a head-final language, meaning that the head follows the complement in the phrase. As illustrated by examples (5a) and (5b), in English the verb precedes the direct object (5a) and in Japanese it is the other way round (5b).

(5a) John [loves Mary]  
     [V O]  
(5b) John-ga [Mary-o butta]  
     John Mary hit  
     [O V]

A child learning English would rely on exemplars like the sentence in (5a) to select the value [VO] of this parameter. The input children receive is, therefore, the trigger for the setting of parameters. Since children’s early productions are consistent with the basic word order of their native language, children must have already fixed this parameter when they begin to produce multiword utterances.

Experimental studies have been conducted to study the timing of the Head direction parameter setting. For instance, Hirsh-Pasek and Golinkoff (1996) observed that, already at 17 months of age, infants use their knowledge about word order to comprehend sentences. Using the intermodal preferential looking paradigm, they presented infants with the sentence “Big Bird is tickling cookie monster” and with two concurrent videos. The video that matched the sentence showed Big Bird tickling Cookie Monster, while non-matching video showed Cookie Monster tickling Big Bird. Infants looked longer to the matching video than to the non-matching one, showing that they comprehend the active reversible sentences.
This evidence supports the hypothesis that early in their life, already before they begin producing multiword utterances, children have fixed the value of the Head direction parameter. How are they able to do this? Infants can set the Head direction parameter by means of the properties of the input they receive. In fact, it has been pointed out the prosodic prominence is a reliable cue to word order. The complement is always the prominent element within the constituent, therefore the location of the prominence varies according to the language typology. Head-initial languages are characterized by prominence at the end of the constituent, while head-final languages are characterized by prominence at the beginning of the constituent (Nespor & Vogel, 1986). Using the high-amplitude sucking paradigm, Christophe and co-workers (Christophe, Nespor, Guasti, & Van Ooyen, 2003) investigated whether infants as young as 6-12 weeks were able to discriminate two languages that differ in their head direction and in its prosodic correlate (i.e., French and Turkish). Original sentences produced by native speakers of those languages were resynthesized so that they contained only Dutch phonemes to make sure that speakers were not classifying languages on the basis of phonemic properties. Results showed that infants were indeed able to discriminate the two languages. The authors stated that the most likely interpretation for their results was that infants relied on the prosodic difference linked to the prominence to discriminate languages. Therefore, their observation supported the hypothesis according to which infants can rely on prosodic information to set the Head direction parameter (see also Gervain & Mehler, 2010; Gervain, Nespor, Mazuka, Horie, & Mehler, 2008; Guasti, Nespor, Christophe, & Van Ooyen, 2001 for further evidence about word order prelexical representation).

In summary, according to this theory infants are endowed with a system of principles and parameters. In order to acquire their language, they have to select one of the two values of a parameter relying on their experience. This approach offers a plausible model of the acquisition of syntax because it introduces a grammar with a limited number of rules. This approach accounts for the fact that children are endowed with linguistic principles that prevent certain operations or organization of words from being possible, therefore they consider only a limited set of all possible rules that could be induced from the input. As Pinker (1994) stated:
“If this theory of language learning is true, it would help solve the mystery of how children’s grammar explodes into adult-like complexity in so short a time. They are not acquiring dozens or hundreds of rules; they are just setting a few mental switches.”

1.2.1.2. Usage-based Theory

An alternative framework suggests that infants’ linguistic knowledge might be based on imitation and item-based constructions. It has been proposed that children move gradually from a concrete knowledge of words, expressions and substantive constructions to the abstract schematic constructions proper of adult grammar knowledge. Therefore, children would not be endowed with an innate abstract linguistic knowledge. On the contrary, as Tomasello (2000, p. 156) stated, “children imitatively learn concrete linguistic expressions from the language they hear around them”.

The approach is known as Usage-based theory (Tomasello, 2000; Tomasello & Brooks, 1998). The theory is grounded in the constructivist approach, which considers adult grammar as an inventory of constructions2 (Goldberg, 1995). The general characterization of language acquisition proposed by Tomasello (2003), and clearly summarized by Ambridge (2004), is the following:

a) The child uses intentional-reading skills (joint attention and cultural learning) to learn a number of utterances, their communicative function, and meaning.

b) The child decomposes these utterances into component parts and generalizes across utterances with shared lexical items (e.g., Where’s Mummy? Where’s Daddy? Where’s X?).

c) The child detects the analogies between these partially abstract constructions and moves to more abstract constructions (e.g., SVO transitive construction).

---

2 Goldberg (1995, p. 4) defines constructions as “the basic units of language. Phrasal patterns are considered a construction if something about their form or their meaning is not strictly predictable from the properties of their component parts or from other constructions.” In adult grammar, constructions are characterized with different levels of abstraction ranging from substantive constructions (e.g., He’s [VERB]-ing) to schematic constructions (e.g., [SUBJECT] [VERB] [OBJECT]).
d) The child restricts the usage of certain lexical items to particular constructions in which they are considered grammatical. For instance, if a verb is presented always in the same argument structure construction such as *sleep* in *the boy sleeps, the rabbit sleeps, the teddy bear sleeps*, then the child infers that the use of that verb in a non-attested construction as, for instance, the transitive one, is not acceptable.

e) At the same time as a), b), c) and d), the child uses the information resulting from the distributional analysis of the input to form the syntactic categories (e.g., VERB, NOUN, etc.).

In the present section we will focus on two issues which have been extensively explored in the context of Usage-based theories, namely *verbs’ argument structure acquisition* and *word order representation*. A series of observational studies of children learning English and other languages (see for example Behrens, 1998 for Dutch, German and English; Pizzuto & Caselli, 1992 for Italian) highlighted item-based patterns (i.e., lexical specific constructions) especially in the use of verbs. This concept has been further developed through experimental studies implying teaching novel verbs to children. These studies are based on the following rationale: if children use novel verbs in a creative way (i.e., in a construction different from the one used by the experimenter in the training phase), it means that they have acquired an abstract schema; however, if they use those verbs in the same way (i.e., in the same construction) they have heard from the experimenter, it means that they their knowledge is linked to a particular verb. For instance, Tomasello and Brooks (1998) studied the performance of 2 and 3-year-old children in the novel verb paradigm. They taught children two novel verbs, one verb was always presented in an intransitive construction while the other verb was presented in a transitive one, In the test phase children were questioned in order to elicit novel uses of the same verbs. Their results showed that only few children used the verbs in novel constructions (e.g., transitive construction for the verb presented in the intransitive construction). Therefore they concluded that children’s verb knowledge was lexical specific. A similar study explored the learning of passive and active constructions reaching the same conclusions (P. J. Brooks & Tomasello, 1999). For the verb trained in the active construction, only 12% of children produced at least one passive utterance and a few more children (especially 3-year-old children) produced active utterances for the passive trained verb.
Another experimental paradigm (weird word order paradigm) has been used in several studies to investigate children’s word order representation (e.g., Abbot-Smith, Lieven, & Tomasello, 2001; Akhtar, 1999; Matthews, Lieven, Theakston, & Tomasello, 2005). This methodology implies children familiarization with novel verbs introduced in non-canonical word order sentences (e.g., Subject Object Verb (SOV) in English). It investigates whether children use the novel verb in a sentence following the example of adults’ production or choosing the canonical order of their native language (i.e., Subject Verb Object (SVO) in English). For example, Akhtar (1999) taught children aged between 2 and 5, three novel verbs each presented in a different sentence structure. One verb was introduced in a SVO structure (e.g., Big Bird dacking the car), another verb in a SOV structure (e.g., Big Bird the car gapping), and another in a VSO structure (e.g., Tamming Big Bird the car). Subsequently, children were asked to describe the action using the novel verb. Results showed that older children used the canonical English order SVO in each condition (“correcting” the SOV or OVS orders presented during training to SVO). On the contrary, younger children were inclined to replicate the type of word order they had heard from the experimenter. As the authors stated, these results suggested that the acquisition of the canonical word order construction is a gradual process. Initially young children’s grammatical knowledge is organized around lexical specific constructions and subsequently these constructions are generalized to more abstract ones. The interpretation of these results has been questioned by Franck and Lassotta (In Press). In their critical review, the authors addressed some theoretical and methodological issues of the weird word order paradigm. They suggested that the data collected, instead of supporting the claims of the usage-based approach, more likely supported the theoretical account according to which from early on the infant represents the word order in an abstract form.

To sum up, the Usage-based approach suggests that children may not be endowed with innate abstract linguistic knowledge. Instead, during the first years of life they might move gradually from item-based constructions to abstract constructions. Initially children rely on imitation of the more concrete constructions (e.g., He’s X-ing) and subsequently, and only in a later phase of development, they develop more abstract schemas (e.g., SVO transitive construction) on the basis of their ability to track similarities.
In the previous sections we outlined the general ideas proposed by contrasting theories accounting for language acquisition in order to present the theoretical framework of this work. In the following sections, we will focus on another issue constituting the theoretical background of this dissertation, namely syntactic processing in natural language. In particular, we will use the case of syntactic ambiguities resolution to sketch the models accounting for syntactic processing in natural language and we will discuss the influence of different factors (i.e., syntactic structure, prosody and semantic) on this process.

1.3. Syntactic Processing in Natural Language

In the following section we will focus on a fundamental process, known as syntactic analysis or parsing. This process allows sentence comprehension describing how the words are structured in a hierarchical representation according to the grammar of a certain language. The interpretation of a sentence is indeed determined by the way the words are combined with each other. As shown in example (6), a sentence may have two different meanings depending on how the words are hierarchically structured.

(6) The girl hit the man with the umbrella.

The girl [[hit [the man]] [with the umbrella]]

The girl [hit [the man [with the umbrella]]].

In one interpretation the umbrella is used as an instrument while in the other the umbrella is a property of the man.

Boland and Blodgett (2001) provide a clear schematization of syntactic parsing, which is described as constituted by three components:

1. *Generation* of the syntactic structure (or of multiple alternative structures, see Interactive accounts)

2. *Selection* of a syntactic structure (if more than one)

---

3 Bracketing indicates the different structures which determine the two possible interpretations of the sentence.
3. Reanalysis if the selected structure results to be incorrect

These three components are not necessarily implemented in different stages in all parsing theories. In parsing theories assuming that the syntactic structure originally generated is only one, such as in the garden path model (Frazier, 1987a), there is no distinction between the first two components. Usually this process is performed quickly and in an accurate way; however, sometimes people may experience some difficulties. This is the case when we process ambiguous sentences.

One type of ambiguity is the so-called syntactic ambiguity. Syntactic ambiguities refer to the fact that more than one syntactic structure could be associated to a sentence and then more than one interpretation would be available. In order to better illustrate this phenomenon, consider the example proposed by Crocker (1999).

(7) John saw the astronomer with the telescope.

In this sentence, with the telescope may be attached to the astronomer, because it is plausible that an astronomer owns a telescope. However, it is possible that with the telescope modifies the verb phrase and it would be interpreted as an instrument of saw. These sentences are called globally ambiguous because both syntactic structures are plausible.

A further kind of syntactic ambiguity, often referred to as the garden-path phenomenon (Bever, 1970), is characterized by a local and temporary ambiguity during the processing of a sentence. It is mainly in these cases that people encounter processing difficulties. This phenomenon is directly linked to the assumption that sentence processing is incremental, since the words are incorporated in the structure as they are encountered and not at the end of the sentence (e.g., Marslen-Wilson, 1973, 1975). Even if this assumption is fairly accepted (Marslen-Wilson & Tyler, 1980; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Traxler, Bybee, & Pickering, 1997), there is no consensus about the stages when different sources of information are taken into account during parsing.

Two main approaches have been proposed in order to characterize sentence processing: modular accounts and interactive accounts (see Mitchell, 1994 for a review).
Modular accounts – These models are based on the assumption that the mind is composed by different modules (Fodor, 1983), which perform separate and specific processes. The general idea shared by these models is that the language processor is composed by several sub-systems which convey the information following a specific direction (serial processing). Such accounts are modular since the processor is assumed to exploit mainly (or only) syntactic information in the initial stages of parsing. Only a single analysis at a time is adopted by the processor and therefore only a single syntactic structure is activated at a time. The best-known modular model is the abovementioned garden-path model proposed by Frazier and her collaborators (Frazier, 1987a). We will develop the discussion of these accounts in the next section.

Interactive accounts – These models are based on the assumption that multiple analyses are performed by the processor and, therefore, more than one syntactic structure is activated at a time; all potentially relevant sources of information can be used immediately during sentence processing (MacDonald, 1993; McClelland, John, & Taraban, 1989). The analysis that receives the most support from varied types of information is the most activated. The difficulty experienced in the processing of ambiguous sentences is due to the competition between several alternatives; when two alternatives have more or less the same level of activation, the processing is more complex.

In the next sections we will sketch the view of modular accounts considering the role of the syntactic structure in parsing.

1.3.1. The Role of Syntactic Structure in Syntactic Processing

In the present paragraph, we will discuss syntactic processing considering the modular accounts. These accounts assume that initial parsing uses mainly (or only) syntactic information. They assume that two principles guide the syntactic processor during parsing (and ambiguity resolution): Minimal Attachment (MA) and Late Closure (LC) (Frazier, 1987a). According to MA, the processor incorporates the incoming material into the tree structure using the fewest number of nodes. This principle explains people’s preferences for a certain type of interpretation (e.g., for the example The farmer hit the tramp with the stick, the
verb-phrase analysis is supposed to be easier than the noun-phrase analysis). When two analyses imply the same number of nodes, the processor applies the LC principle, which predicts that the new material is attached to the currently processed phrase (e.g., Traxler, Pickering, & Clifton, 1998). For instance, in a sentence like (8) it will predict that the clause \textit{that was dark} attaches to the most recent clause \textit{the chocolate} rather than to the higher clause \textit{the cake}.

\begin{itemize}
\item[(8)] The cake with the chocolate that was dark was not eaten by the kids.
\end{itemize}

These principles are based exclusively on syntactic information and do not refer to any other source of information (e.g., lexical frequency, plausibility, etc.) (Ferreira & Clifton, 1986; Ferreira & Henderson, 1990; Rayner, Carlson, & Frazier, 1983). According to the garden-path model these other sources of information can affect later stages in the processing (e.g., re-analysis), but not the initial analysis. Another principle was subsequently introduced, namely the active filler strategy (Frazier, 1987b) or the minimal chain principle (De Vincenzi, 1991). This principle was proposed to account for long distance dependencies such as relative clauses or \textit{wh}-questions, as in (9). In these sentences the argument of the verb (the \textit{filler}) moves from the canonical position after the verb to the beginning of the sentence, leaving a \textit{gap} at its original position.

\begin{itemize}
\item[(9)] Which man do you think Mary loves \textit{\underline{___}} a lot?
\end{itemize}

According to this principle, the first possible gap location is considered to be the correct one until further evidence requires a revision of this interpretation. Therefore in (9), the parser considers first the interpretation in which \textit{which man} is the object of \textit{think} (the gap is identified after \textit{think}). However, when \textit{Mary} is analyzed this interpretation is no longer possible, revision is required and another interpretation has to be considered (the gap is identified after \textit{loves}). Empirical evidence in favour of the existence of this principle comes from the “filler gap” effect. For example, using a self-paced reading task, Stowe (1986) observed slower reading times for the direct object position \textit{us} in sentences like (10a) than for control sentences like (10b). These results suggest that the parser positioned the object gap in the first possible gap location without checking whether the direct object position was occupied.
(10a) My brother wanted to know who Ruth will bring us home to ___ at Christmas.

(10b) My brother wanted to know if Ruth will bring us home to Mom at Christmas.

Although experimental evidence of the validity of these principles has been collected, alternative accounts have been proposed. In these accounts, other aspects of language introducing different sources of information are considered to play an immediate role in syntactic processing. In the next section we will consider the contribution of semantic information to the early stages of syntactic processing as proposed by *Interactive models*.

### 1.3.2. The Role of Semantics in Syntactic Processing

The role of semantics in syntactic processing has been mainly acknowledged by *Interactive models* of syntactic ambiguity resolution (e.g., MacDonald, 1993; MacDonald, Pearlman, & Seidenberg, 1994; Snedeker & Trueswell, 2004; Spivey-Knowlton, Trueswell, & Tanenhaus, 1993; Trueswell, Tanenhaus, & Garnsey, 1994). The general statement of this approach is that syntactic parsing is influenced by all available information types from the very first stages of processing. The typical study used to investigate this statement exploited temporary syntactic ambiguous sentences in which a group of words could be either a modifier or it could introduce another entity. To illustrate this concept we consider in (11) an example proposed by Snedeker and Trueswell (2004).

(11) The chef poked the pastry with the…

The prepositional phrase starting with *with the* could be linked either to the verb indicating an instrument (e.g., *with the fork*) or to a noun phrase indicating a modifier (e.g., *with the flaky crust*). In order to solve this ambiguity, listeners could rely on the referential context of the sentence or on semantic information.

Initially, it has been proposed that the choice of the appropriate syntactic analysis was made on the basis of its accordance with the conversational context (Crain & Steedman, 1985). Considering the previous example, this account predicts that, if two possible referents (two or more pastries) are available in the context, the listeners would prefer the interpretation
involving the modifier; if only one referent is available in the context, the instrument interpretation would be preferred. This assumption has been supported by studies mainly using reading tasks (e.g., Altmann & Steedman, 1988; Hagoort, 2003; Snedeker & Trueswell, 2004; Van Berkum, Brown, & Hagoort, 1999 among many others).

Lexicalist theories of parsing (MacDonald et al., 1994; Trueswell, 1996; Trueswell et al., 1994) subsequently claimed that lexically specific semantic information plays an important and immediate role in syntactic processing (see for example Trueswell et al., 1994; Trueswell, Tanenhaus, & Kello, 1993 among others). For instance, Trueswell and colleagues (Trueswell et al., 1994) investigated the processing of reduced relative clauses. The ambiguity of these sentences arises since the form searched could be either a past tense or a past participle. In the first case the initial noun is the subject and the agent of the verb, while in the second case it is the object and the theme of the verb. The authors observed that the ambiguity was reduced when the initial noun was a poor agent of the verb, like room in (12a), compared to the situation in which the initial noun was a good agent of the verb, like thief in (12b). They stated that this example showed how the semantic fit between the phrases and the possible thematic role influenced and constrained the syntactic processing.

(12a) The room searched...

(12b) The thief searched...

All in all, the crucial question underlying the characterization of syntactic processing is not whether different information types influence processing rather when these different information types influence it. If, on the one hand, there are modular models assuming that it is only in a second processing stage that non syntactic information plays a role, on the other hand interactive accounts acknowledge that referential and semantic information play an immediate role in solving ambiguities by interfering or favouring a syntactic analysis over another. To date, this debate is still ongoing, as empirical findings supporting one or the other approach are not conclusive.

Interactive models are usually characterized as lexicalist in nature, that is syntactic structures are supposed to be generated via the combination of partial structures that are lexically stored and associated with individual lexical items.
1.3.3. The Role of Prosody in Syntactic Processing

The role of another type of information, namely prosody, has been investigated in the context of syntactic processing. Also in this case, although most psycholinguistic approaches acknowledge that prosody is used for syntactic disambiguation and to guide parsing decisions (e.g., Marslen-Wilson, Tyler, Warren, Grenier, & Lee, 1992; Schafer, Carlson, Clifton, & Frazier, 2000; Steinhauer, Alter, & Friederici, 1999), it is still a matter of debate at which stage it intervenes in the processing.

Prosody is generally defined as the sound pattern characterizing any spoken sentence. In particular, as Ferreira pointed out (Ferreira, 2003), prosody is constituted by three main aspects: intonation, timing, and stress. Intonation concerns the variation of the pitch and is characterized as the melody of the utterance. Words falling into the same pitch group are considered as units of the same "intonational phrase". Timing is determined by the duration of the words and the duration of the pauses in the sentence. This aspect is particularly important for syntactic processing since its components (i.e., phrase lengthening and pauses) tend to mark syntactic boundaries in speech. Finally, stress is an aspect of prosody that refers to the difference in loudness between the syllables occurring in a word. The function of prosody we are interested the most in is the disambiguation function. This function assists the syntactic and semantic processing of the sentence, contributing to remove ambiguities.

It has been claimed that prosody and syntax are tightly linked (Nespor & Vogel, 1986) and it has been shown that prosodic representations influence the production and the comprehension of the syntactic structure of the sentences. Therefore, it is of interest to explore how these two aspects of language interact. As previously mentioned, experimental evidence concerning the interaction between prosody and syntax has been mainly collected in studies involving some kind of syntactic ambiguity. In fact, as Nicol (1996) stated, when a sentence is not syntactically ambiguous, prosody adds little information to their comprehension; however, when the sentence is ambiguous, prosody may be used to guide syntactic parsing. In order to illustrate this statement we will consider the same example used by Nicol and presented (13a).

(13a) When you learn gradually you worry more.
The two possible interpretations are illustrated in (13b) and (13c) in which the comma indicates the prosodic boundary:

(13b) When you learn, gradually you worry more.

(13c) When you learn gradually, you worry more.

We observe that the presence of prosodic cues characterizing prosodic boundaries (pause and phrase lengthening) indicates where syntactic units end and, consequently, where the listener has to put a syntactic boundary. This fact results in two different meanings of the sentence.

The assumption drawn from this example has been demonstrated experimentally in a series of studies showing that speakers produce prosodic cues to disambiguate sentences (see for example Kjelgaard & Speer, 1999; Marslen-Wilson et al., 1992; Schepman & Rodway, 2000 among others) and also that adults use this kind of cues in sentence comprehension. For instance, Kjelgaard and Speer (1999) observed that their participants (adult English speakers) benefited from the presence of prosodic boundaries coinciding with syntactic ones in a task involving ambiguous sentences. In a speeded grammaticality judgment task\(^5\), prosodic structure (the presence of phrase lengthening and pauses) influenced positively participants’ performances: fewer errors and shorter reaction times were observed compared to other conditions (baseline and contrasting prosody). These results were replicated with other tasks, such as sentence comprehension and sentence completion. Therefore, it has been claimed that prosody has early effects on sentence comprehension, facilitating or interfering with syntactic processing (see for example Kang & Speer, 2004; Watson & Gibson, 2004, 2005 among others). Empirical evidence of the early influence of prosody in the initial stages of syntactic parsing also comes from ERP investigation. For example, Eckstein and Friederici (2006) observed that when participants presented with combined syntactic and prosodic violations, an early temporal negativity was present bilaterally.

To sum up, in this section we presented evidence about the influence of prosodic information on syntactic processing. In particular, we highlighted the relation between

---

\(^5\) Participants were instructed to answer as quickly as possible.
prosodic and syntactic boundaries. We observed that, although this relation is not exactly symmetrical (not all the syntactic boundaries are marked with prosodic boundaries, although all the prosodic boundaries correspond to syntactic boundaries), it represent a useful tool in the hands of the listener to interpret and to disambiguate sentences.

1.4. Artificial Language Learning Paradigm

In order to acquire more information about the mechanisms that enable natural language learning, scientists have looked more and more at artificial language learning (ALL) paradigms. In ALL paradigms, participants are exposed to structured materials following certain rules. These rules mimic aspects of natural language structures but are simplified relative to the realistic language learning conditions. These controlled stimuli are presented to participants during a training phase, subsequently participants are tested in order to study their ability to extract and generalize those regularities. The principle behind this kind of studies is to use simplified systems to investigate different aspects of processing independently of one another. Initially, as deVries (2011) pointed out, ALL paradigms were used to study implicit sequence learning. A rationale for this statement can be found in Conway and Pisoni (2008), who suggested that language acquisition and implicit sequence learning share at least two aspects. First, both processes require the extraction of sequential patterns from a complex input; second, both processes are characterized by the difficulty to explicitly verbalize what has been learned. Subsequently, ALL paradigms have been employed to investigate different aspects of natural language learning in isolation. Gomez and Gerken (2000) identified four main topics investigated with ALL paradigms: word segmentation, word order, grammatical relations (which correspond to artificial grammar learning studies), and syntactic categories.

One of the first difficulties infants have to face in learning language is how to segment words from the continuous speech stream they experience (see Section 1.4.2.1.1.). Studies have shown that infants possess a sophisticated mechanism that enables them to segment the speech based on statistical computations (e.g., Saffran, Aslin, & Newport, 1996; Saffran, Newport, & Aslin, 1996). These seminal studies were followed by others showing that this statistical learning mechanism can be exploited also with other domains stimuli such as the
visual or the musical ones to identify units from continuous streams (e.g., Fiser & Aslin, 2002a; Gebhart, Newport, & Aslin, 2009; Saffran, Johnson, Aslin, & Newport, 1999).

Infants not only learn how to segment words, but also how to put them into the right order in sentences (see Section 1.4.2.1.2.). ALL experiments have been used to determine whether infants were able to discriminate sentences generated by different grammars based on the different order of the words. These studies suggested that indeed infants are able to do so in a quick and accurate way (Gomez & Gerken, 1999). Other research showed that infants are not only able to acquire the order of the words in sentences they are also able to extract more abstract patterns from specific word order. Very early in their life they are able to generalize their knowledge of abstract patterns over novel situations (e.g., novel lexicon) (Marcus, Vijayan, Bandi Rao, & Vishton, 1999).

Finally, the last topic explored in ALL experiments concerns syntactic categories learning (see Section 1.4.2.4.). The ability to abstract over categories is very important for language acquisition. In fact, if we recognize that a novel word belongs to a certain syntactic category we automatically know that all the properties that are valid for that category apply to the novel word (e.g., if an English speaker recognizes that wug is a noun, she should be able to create the plural form wugs even if she never heard that expression before). This type of abstraction has been investigated with ALL paradigms in which participants had to learn to categorize novel words and then discover the relationships between these categories. Interestingly several experiments showed that adults are able to categorize novel words based on the perceptual cues that indicate category membership (e.g., Frigo & McDonald, 1998).

There are several advantages in using ALL paradigms for the broader study of language learning. Gomez and Gerken (2000) emphasized two of them: 1) ALL studies allow the experimenter to control the input to which the participants are exposed; 2) ALL studies allow the experimenter to control for prior learning. Controlling the input that learners receive permits the study of aspects of language in isolation. Because natural language is a complex system, it is difficult to disentangle the several aspects playing a role during acquisition. Therefore, the use of artificial languages allows researchers to avoid redundant and/or confounding features which may influence the learning process. Notwithstanding the previous
advantages, ALL paradigms should always be confronted with the realistic conditions of language acquisition in order to enhance their validity.

In the following sections we will outline the methodology used in ALL studies and we will perform a review of the literature concerning ALL from several points of view. In the methodology section (Section 1.4.1.), we will address general topics concerning the way ALL studies are implemented. Subsequently, in the literature review (Section 1.4.2.) we will summarize the main results of experiments in ALL.

1.4.1. Methodology

In the next sections we will introduce some methodological aspects of ALL experiments. Firstly we will outline the characteristics of the artificial languages created for these studies; in particular, the types of units employed and the types of dependencies studied will be presented. Subsequently, the different types of experimental tasks will be summarized with particular attention to the types of instruction given to the participants and the types of task used. Finally, we will present the populations whose performance has been investigated in ALL studies.

1.4.1.1. Artificial Languages

Different artificial languages have been used in ALL experiments. These languages can vary along several dimensions. In the following sections we will mainly overview the types of patterns and dependencies investigated, which characterize the properties of artificial languages, and the types of units used (i.e., linguistic or non-linguistic).

Properties of Artificial Languages

Artificial languages are characterized by several properties. In this section we will present the most common patterns of regularities investigated in ALL experiments. According to the aim of the research, different types of dependencies have been used to generate the target language of the experiment. Some of these regularities are defined as particular patterns, while others are only classified as adjacent or non-adjacent dependencies. Examples of the first type of regularities can be primarily found in experiments exploring syntactic structures
learning (see Section 1.4.2.1.2.). For instance, we can identify finite state grammars and phrase structure grammars (e.g., Bahlmann, Gunter, & Friederici, 2006; Bahlmann, Schubotz, & Friederici, 2008; Fitch & Hauser, 2004; Reber, 1967). AXC-languages (e.g., Gomez, 2002; Peña, Bonatti, Nespor, & Mehler, 2002), and ABA-ABB patterns (Marcus et al., 1999). When the experiment focused on the segmentation problem (see Section 1.4.2.1.1.), the regularities patterns generating the artificial language could be usually classified in one of the following two groups: adjacent dependencies or non-adjacent dependencies.

Initially, ALL experiments focused on the relationships between adjacent elements and showed that participants were able to use transitional probabilities available in the input to track this kind of relationships (e.g., Aslin, Saffran, & Newport, 1998; Hauser, Newport, & Aslin, 2001; Reber, 1967; Reber, Kassin, Lewis, & Cantor, 1980; Saffran, Aslin, et al., 1996; Saffran et al., 1999; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997). However, natural languages also include non-adjacent dependencies between units. For instance, it is possible to find relationships between non-adjacent words in the syntactic structures (e.g., materials may intervene between subject and verb in number agreement like in the sentence The daughters of the neighbour are hungry), as well as between sounds within words (e.g., in Semitic languages vowel patterns inserted between the consonants in a word determine number or signal time). Therefore, researchers extended their investigation to the learnability of non-adjacent dependencies in ALL paradigms (Bonatti, Peña, Nespor, & Mehler, 2005; Creel, Newport, & Aslin, 2004; Gomez, 2002; Misyak & Christiansen, 2007; Misyak, Christiansen, & Tomblin, 2009; Newport & Aslin, 2004; Newport, Hauser, Spaepen, & Aslin, 2004; Onnis, Christiansen, Chater, & Gomez, 2003; Perruchet & Rey, 2005; Perruchet, Tyler, Galland, & Peereman, 2004; Toro, Nespor, Mehler, & Bonatti, 2008). These studied showed that non-adjacent dependencies learning is more complex than adjacent dependencies learning and requires particular circumstances such as highly variable context and/or the occurrence of perceptual cues.

Types of Units

Several types of units have been employed to study ALL (e.g., syllables, phonemes, tones, geometrical shapes, etc.). These units can be categorized into two groups: linguistic
materials and non-linguistic materials. The linguistic materials group is further constituted by various subcategories of units such as phonemes, syllables, and words (see also Section 1.4.2.1.). Other subcategories such as tones, geometrical shapes, and tactile stimuli can be found in the non-linguistic materials group (see also Section 1.4.2.2.).

**Linguistic materials:** We will consider now two main types of linguistic materials, phonemes (consonants and vowels) and syllables, that have been exploited in ALL experiments. The use of different types of linguistic materials allowed experimenters to study the constraints and limitations in regularities learning. The first studies in ALL used sequences of consonants generated by finite state grammars to study adults’ capacity to learn implicit grammatical patterns (e.g., Reber, 1967, 1969). Subsequently, recent studies showed a different function for vowels and consonants in a task requiring regularities extraction (Bonatti et al., 2005; Newport & Aslin, 2004; Toro et al., 2008). These studies showed that the information carried by vowels was employed to extract structural generalizations while the computation of transitional probabilities over consonants played a role in word identification.

The great majority of the ALL studies used syllables. Several studies explored the capacity of adults, infants and animals to track transitional probabilities over syllables in order to segment continuous speech (e.g., Aslin et al., 1998; Peña et al., 2002; Saffran, Aslin, et al., 1996; Saffran, Newport, et al., 1996, 1997; Shukla, Nespor, & Mehler, 2007 among others), while other studies used syllables as (monosyllabic) pseudo-words to study the learning of grammatical patterns (e.g., Braine, 1963, 1966; Gomez & Gerken, 1999; Marcus et al., 1999; Saffran et al., 2008 among others). This last kind of study was also performed with neuroimaging techniques in order to detect the neural networks activated by artificial language processing (Bahlmann et al., 2006, 2008; Friederici, Bahlmann, Heim, Schubotz, & Anwander, 2006; Mueller, Bahlmann, & Friederici, 2010).

---

6 Transitional probability of an element Y given an element X is computed as follows:

When the transitional probability is high, the presence of X strongly predicts the occurrence of Y; when the transitional probability is low, the co-occurrence of X and Y is weak.
**Non-linguistic materials:** It has been shown that the capacity to track regularities patterns between adjacent and non-adjacent elements is not limited to linguistic sounds but can be applied to non-linguistic sounds as well (Creel et al., 2004; Endress, 2010; Gebhart et al., 2009; Saffran et al., 1999). In addition, regularities extraction has been investigated using visual stimuli, such as sequences of geometrical shapes (Fiser & Aslin, 2001, 2002b, 2005; S. P. Johnson et al., 2009; Tettamanti et al., 2009), tactile stimuli (Conway & Christiansen, 2005), and also as visuomotor learning employing the serial reaction time task (Hunt & Aslin, 2001). The use of non-linguistic materials aimed to explore whether the mechanisms subserving learning were domain-specific or domain-general and whether across domain generalization was possible (see Section 1.4.2.4. for further details about generalization).

### 1.4.1.2. Experimental tasks

In the present section we will introduce how ALL has been investigated. Since different experimental tasks have been exploited to study ALL, we will outline these procedures according to two topics: types of instruction and types of task.

#### Types of instruction

The type of instruction (if any) given to participants before the training phase influences the learning processing performed by participants themselves (Reber et al., 1980). In some experiments participants are explicitly asked to memorize the sentences presented during the training phase (e.g., Knowlton & Squire, 1996; Poletiek, 2002; Reber, 1967, 1969). In other experiments participants are asked to search for the rule(s) characterizing the items of the training phase (e.g., Mathews et al., 1989; McAndrews & Moscovitch, 1985; Tettamanti et al., 2002) or are directly told the rule (Musso et al., 2003; e.g., de Graaff, 1997). In contrast, in some ALL experiments no explicit instructions are given; participants have only to passively listen to (or watch) the stimuli presented during the training phase (e.g., Saffran et al., 1999; Saffran, Newport, et al., 1996, 1997; Zimmerer, Cowell, & Varley, 2010).

Indeed this variability influences the type of learning that participants are supposed to do. For instance, when there are no instructions (or the instructions do not directly refer to regularities extraction), it is highly possible that participants perform implicit learning. On the
other hand, when they are instructed to look for regularities or are directly told the rule, other types of learning, such as explicit learning are explored. The effect of the type of instruction on artificial language learning has not been systematically investigated. Initially, it was suggested that instructions leading to implicit learning would have favoured the learning of complex structures (e.g., Reber et al., 1980). More recently the few studies comparing the learning in the implicit and in the explicit instruction conditions suggested an advantage for the latter (de Graaff, 1997).

**Types of the task**

Different tasks are employed in ALL experiments; the most common ones are listed below.

*Word recognition task:* this task is constituted by a training phase in which participants are passively exposed to a continuous speech stream and subsequently they are tested on word recognition. In the test phase participants have to choose between a word which was actually presented during the training phase and another word which could be either a part-word or a rule-word. Part-words are the merge of parts of two words presented during the training. For instance, if in the training the words [puliki] and [taRadu] were presented, a part-word could be [kitaRa], which is the merge of the last syllable of a word and the first two syllables of the other one. Rule-words are words that follow the same pattern of the words presented during training but they are built up with different units, for example if in the training phase the words [puliki] and [puRaki] were presented, a possible rule-word could be [pubeki] (e.g., Bonatti et al., 2005; Gomez, 2002; Peña et al., 2002; Saffran, Newport, et al., 1996; Toro et al., 2008). The number of correct responses (mean score) is calculated and compared to a performance at chance (50% correct responses) with single-sample t-tests. It is possible to claim that participants are able to segment the signal and successfully recognize the words when participants performance is significantly better than chance (see Section 1.4.2.1.1. for further details about the experiments).

*Familiarization preference task:* this task is used with infants. In this task infants are sitting on their caregiver’s lap in front of a video monitor. Two lights are positioned on the sides with speakers beneath them. The auditory material is played from the speakers. This procedure is
constituted by two phases: 1) a training phase in which infants are exposed to auditory materials which can be constituted by a continuous stream of sounds or by a list of sentences separated by pauses; 2) a test phase in which two types of stimuli are presented, namely stimuli that were contained within the familiarization materials (consistent items) or stimuli similar but that were not contained in the familiarization materials (inconsistent items). In the test phase infants can control the duration of the presentation of the sound files as sounds start playing when the infant looks towards the blinking light and stop when the infant looks away for more than 2 seconds (e.g., Marcus et al., 1999; Saffran, Aslin, et al., 1996; Saffran et al., 2008). The dependent variable measured in this type of task is the mean listening time, that is the amount of time infants spend to listen to the consistent or to the inconsistent items. A significant difference between the listening times for the two types of items is interpreted as the fact that infants recognized the words with which they have been familiarized. Infants usually prefer to listen to the inconsistent items since they represent a novel, more interesting input.

**Grammaticality judgment task:** this task is also constituted by two phases. In the familiarization phase participants listen to lists of sequences generated by an artificial language and are instructed to pay attention to the stimuli since they will be tested on them later. After the familiarization, participants are informed that the sequences they heard were characterized by grammatical regularities. In the test phase participants are asked to judge whether the stimuli they listen to follow the same grammar or not (Misyak & Christiansen, 2007; e.g., de Vries, Monaghan, Knecht, & Zwitserlood, 2008). The percentages of correct responses and/or d-prime scores are calculated as dependent variables. Group performance is then compared to a performance at chance (50% or 0 for the d-prime scores) with a single-sample t-test. It is possible to claim that participants are able to segment the signal and successfully recognize the words when participants’ performance is significantly better than chance (see Section 1.4.2.1.2. for further details about the experiments). It is of interest to mention that in some studies using the grammaticality judgment task, participants received feedback (e.g., Bahlmann et al., 2006, 2008; Friederici et al., 2006; de Vries et al., 2008). Since the role of feedback in language acquisition is highly controversial (see Valian, 1999 for a review), little attention has been devoted to the investigation of its role also in the context of
artificial language learning paradigms. To our knowledge only one study (Dale & Christiansen, 2004) directly investigated the influence of feedback in an ALL experiment. The authors showed that feedback allowed participants to reach a high level of competence in the artificial language. Even if the idea that learners can acquire natural language without any kind of feedback is commonly accepted, further research could highlight its influence in particular aspects of artificial language learning.

**Discrimination task:** this task is used with non-human primates, usually cotton-top tamarin monkeys. It corresponds to the word recognition task used with human adults. The first day, tamarin monkeys are exposed for 21 minutes to a continuous stream of syllables characterized by regularities. On the second day, they are placed individually in a sound-proof chamber and familiarized with two more minutes of the speech stream. Subsequently, they are presented with test trials constituted either by words taken from the familiarization stream or part-words (which were built up by one syllable from a word and two syllables from a different word). The test trial is presented when the monkey faces 180° away from a concealed speaker. The dependent measure is an orienting response towards the stimulus presented from the speaker. As well as infants, non-humans primate that have successfully learned to segment the input during training, usually preferred to look longer towards the speaker playing the inconsistent items since they represented a novel, more interesting input (e.g., Fitch & Hauser, 2004; Hauser et al., 2001; Newport et al., 2004; Toro & Trobalón, 2005).

**1.4.1.3. Population**

We can identify four types of population whose performance has been studied in ALL paradigms: adults, infants, animals, and patients with neurological or speech disorders. At first, artificial language learning experiments were mainly conducted on adults (e.g., Braine, 1963, 1966; Reber, 1967, 1969), but in the last twenty years a growing interest in performing these studies with infants has arisen (e.g., Aslin et al., 1998; Gomez & Gerken, 1999; Marcus et al., 1999; Saffran, Aslin, et al., 1996). The study of infants’ performance to ALL paradigms allows researchers to make inferences about the leaning mechanisms available during infancy and their innateness. Moreover, the possibility to test infants and children at different ages
should result in a better characterization of the developmental trajectories of the cognitive mechanisms subserving language acquisition.

*Patients with neurological disorders* (e.g., Parkinson’s disease, agrammatic aphasia) have been tested on grammatical pattern extraction in artificial language learning experiments (e.g., Christiansen, Louise Kelly, Shillcock, & Greenfield, 2010; Peigneux, Meulemans, Van der Linden, Salmon, & Petit, 1999; Witt, 2002; Zimmerer, 2010). The comparison between the performance of the group of patients with the performance of a matched control group affords researchers with two kinds of information: the existence of domain-specific or domain-general mechanisms for language processing, and evidence about the neural networks underling the processing of structural regularities. For instance, the observation that language impaired patients failed in performing an artificial grammar learning task with non-linguistic items, argued in favour of a domain-general mechanism subserving learning both in a non-linguistic and in language domain (Christiansen et al., 2010).

ALL paradigms have also been used to study the learning abilities of a few *animal* species. The aim of these studies was to assess whether the capacity to segment speech and to extract grammar-like patterns was unique to humans or might be exploited also by other species. Most studies have focused on non-human primates, namely cotton-top tamarins (e.g., Fitch & Hauser, 2004; Hauser et al., 2001; Newport et al., 2004; Saffran et al., 2008). Early studies explored their ability to segment units from a continuous speech computing statistical regularities. The authors of these studies claimed that important commonalities were observed between infants and non-human primates in their capacity to compute transitional probabilities between syllables (Hauser et al., 2001).

Animals were studied also to investigate their capacity to learn grammar-like patterns (Fitch & Hauser, 2004; Gentner, Fenn, Margoliash, & Nusbaum, 2006; Newport et al., 2004; Saffran et al., 2008). Contrasting conclusions about their capacity to learn phrase structure grammars emerged. Some studies showed that non-human primates were able to learn only finite state grammar, while other studies, in which songbirds were tested, showed learning of phrase structure grammar as well.
In sum, ALL experiments are characterized by high variability in the way they are carried out. In particular, several aspects may vary such as the type of artificial language and the experimental task chosen. These aspects can change according to the aim of the study, to the target population, and to the specific aspects of the learning process the researcher wants to uncover. In the next sections, we will summarize the main results of experiments in ALL.

1.4.2. Review of the literature in ALL

In the previous sections we outlined the methodological aspects characterizing ALL experiments; in the following ones we will review the literature concerning ALL. We will begin our review summarizing studies performed in the linguistic domain. Subsequently, we will present experiments performed in non-linguistic domains. As we will show in these sections, learners possess powerful cognitive mechanisms able to track adjacent and non-adjacent dependencies between units of different nature. This capacity needs to be constrained, as it may result in a large number of computations. We will present two types of constraints that may play a role during language acquisition. In the last two sections different types of generalization explored in ALL experiments will be outlined and the main theories accounting for the type of knowledge resulting from ALL will be presented.

1.4.2.1. ALL in the linguistic domain

ALL studies have investigated the presence of at least two distinct learning mechanisms, a statistical learning mechanism devoted to tracking linear order on the basis of transitional probabilities and a mechanism devoted to structure (grammar-like) learning. Since infants and adults are indeed able to perform statistical computations over an artificial input (e.g., Saffran, Aslin, et al., 1996; Saffran et al., 1999; Thompson & Newport, 2007), it has been suggested that this capacity may be powerful enough to subserve the acquisition of language (Seidenberg, 1997).

However, results of other ALL experiments underlined the fact that, in order to acquire language, the presence of a learning mechanism dealing with grammar-like patterns was necessary. For instance, both Marcus and colleagues (1999) and Pena and co-workers (2002) proposed the existence of an additional learning mechanism of an algebraic-like or rule-like
governed nature. The reason of the necessity of this type of mechanism came from the fact that when the test materials differ in some respect (usually the vocabulary) from the training materials, learners could not simply rely on the statistics calculated over the specific items during training. If the lexicon used in the test phase is different from the lexicon used during training it is not possible that the learning of patterns or relationships is based solely on transitional probabilities’ computation. A mechanism able to abstract beyond the specific items is required. ALL studies are often used to collect evidence in favour of one or the other theory accounting for the nature of the learning mechanisms and of the constraints to those mechanisms required for language learning.

While statistical learning mechanisms have been mainly investigated in the context of segmentation studies, the existence of a mechanism devoted to structure learning has been essentially explored in studies concerning syntax learning. In the former participants usually have to extract words (or word-like units) from a continuous stream, while the latter refers to the detection and generalization of syntactic structures from lists of sentences generated from a target artificial grammar.

1.4.2.1.1. Segmentation

During language acquisition, one of the first tasks language learners are faced with is the need to extract words from continuous speech. As a matter of fact words are not usually delimitated by clear cues such as pauses. Several ALL studies, investigating both infants and adults, have shown that learners are able to perform statistical computations, such as calculating transitional probabilities, over units in order to segment the input (see for example Aslin et al., 1998; Saffran, Aslin, et al., 1996; Saffran et al., 1999). Transitional probabilities computation has been investigated both with reference to adjacent dependencies and non-adjacent dependencies.

Adjacent Dependencies

In the first ALL studies focusing on word segmentation, Saffran and colleagues (Saffran, Aslin, et al., 1996; Saffran, Newport, et al., 1996) exposed adults and 8-month-old infants to a spoken nonsense language stream. The only cues to word boundaries were
transitional probabilities among adjacent syllables (see an extract in Figure 1), which were higher between syllables belonging to the same word and lower between syllables belonging to different words. In the test phase, participants were able to discriminate pseudo-words (e.g., *golabu*, *bidaku*) from part-words (e.g., *kupado*, *tigola*). In order to perform correctly this task, participants should have been able to calculate the statistics characterizing the input to find pseudo-words’ boundaries.

\[ \text{...}\text{bidakupadotigolabubidaku...} \]

\[ \text{0.33} \quad \text{1.0} \]

**Figure 1.** Example of continuous stream presented to the participants to Saffran et al. (1996) study, pseudo-words are marked with different colours.

In addition, the role of prosody in speech segmentation has been investigated with ALL experiments. For instance, Bagou, Fougeron and Frauenfelder (2002) explored the contribution to word segmentation of two prosodic cues, namely lengthening and f0 rise on the final syllable, in a miniature language. Adult participants were exposed to a 12 minutes speech stream and subsequently tested in a lexical discrimination task. The comparison between performance in four conditions (i.e., no cues available, only lengthening available, only f0 rise available, and both cues available) showed that the presence of at least one prosodic cue enhanced performance in the task. In particular, f0 rise resulted in a slightly more accurate performance than lengthening. Therefore, prosodic cues seem to facilitate word segmentation from a continuous speech stream.

**Non-Adjacent Dependencies**

According to several studies non-adjacent dependencies learning is more complex than adjacent dependencies learning and requires particular circumstances. For example, Newport and Aslin (2004) observed that the *nature of the units* involved in the dependency plays a role in its learnability. In their first experiment, participants (undergraduate English speakers) were exposed to a continuous stream constituted by trisyllabic (CV-CV-CV) pseudo-words randomly ordered. The transitional probabilities between the adjacent syllables (within and
between words) varied from .20 to .25, while the transitional probability between the first and the third syllables within words was set to 1.0. Therefore, in order to be able to segment the continuous speech, participants had to rely on non-adjacent dependencies regularities. In the test phase, the authors observed that participants failed to discriminate the words of the language from part-words. This observation indicated that participants were not able to learn non-adjacent dependencies among syllables. In other experiments, the authors tested whether participants were able to track regularities between non-adjacent phonemic segments (i.e., consonants in Experiment 2). In this novel artificial language the transitional probabilities between the consonants building up the pseudo-words were set to 1.0, while transitional probabilities between the vowels were lower. Participants were successful in discriminating words over part-words, showing that they were capable of computing the transitional probabilities among non-adjacent consonants. The same positive result was replicated when the transitional probability was equal to 1.0 over vowels but not over consonants. Interestingly, these observations reflect the pattern found in natural languages, such as the Semitic languages, that exploit the relationship between consonants and vowels, but not between syllables, to carry number information.

Subsequently, Bonatti and colleagues (2005) showed functional differences between consonants and vowels performing two experiments similar to those of Newport and Aslin (2004). Their participants (undergraduate French speakers) were presented with continuous speech streams constituted by trisyllabic (CV-CV-CV) pseudo-words. The transitional probabilities between consonants were set to 1.0 (0.5 at word boundaries) while the transitional probabilities between vowels were equal to 0.5. As in the first experiment performed by Newport and Aslin (2004), participants were able to discriminate words over part-words showing that they relied on transitional probabilities among consonants to segment the signal. In the second experiment, Bonatti and colleagues tested whether participants were equally successful when high transitional probabilities were defined over vowels. Therefore, in this experiment the transitional probabilities between vowels were set to 1.0 while the transitional probabilities between consonants were equal to 0.5. Results showed no discrimination between words and part-words, suggesting that participants were not able to segment the signal relying on vowels’ transitional probabilities. The authors suggested that the
difference in the performance between these two experiments could be explained by the fact that, in natural language, consonants seems to preferentially contribute to the lexical processing of the words (Nespor, Peña, & Mehler, 2003). Therefore, it is plausible that consonants are the units used to segment the speech.

To sum up, several studies underlined the fact that learners are equipped with a powerful statistical learning mechanism. This mechanism seems to play an important role in signal segmentation allowing learners to identify units in a continuous speech stream. While adjacent dependencies are readily acquired, tracking non-adjacent dependencies requires specific circumstances. In the context of speech segmentation, the main factor influencing non-adjacent dependencies computations seems to reside in the nature of the units involved in the dependency.

1.4.2.1.2. Syntax Learning

ALL experiments, and especially artificial grammar learning studies, have been used to acquire information about syntax learning in infants and adults. Indeed, once learners have acquired the words of the lexicon they have to acquire structure regularities generated by grammars. The learning of both adjacent and non-adjacent syntactic dependencies has been explored.

Adjacent Dependencies

Early studies (see for example Reber, 1967; Reber, 1969 among others) in ALL focused on the ability of participants to discriminate grammatical and ungrammatical letters strings generated by a finite state grammar (FSG). Finite State Grammars are characterized by adjacent dependencies since they can be specified by the transitional probabilities between the elements in the sequence. In his influential work, Reber (1967) asked participants to perform multiple learning sessions in which they had to memorize strings of letters which were generated by a FSG like the one presented in Figure 2. Subsequently, in the test phases they were asked to write the strings they had memorized on a piece of paper. The analysis of mean errors showed that they became increasingly sensitive to the underlying structure of the stimuli, even if they were not aware of the rule characterizing the grammatical items.
Chapter 1. Theoretical Framework

Figure 2. Example of Finite-State Grammar.

Gomez and Gerken (1999) tested the capacity of 12-month-old infants to extract the regularities generated by a finite state grammar using the head-turn preference procedure. Infants were trained with sequences of monosyllabic pseudo-words generated by a FSG. The authors tested their ability to discriminate between new grammatical and ungrammatical sequences in four different experiments. In the first experiment infants were tested with ungrammatical sentences characterized by illegal endpoints. For example, during the training phase they were exposed to a finite state grammar in which the first word of the strings could be either VOT or PEL and the last word could be either TAM or JIC. In the test phase they were presented with novel sentences (generated by the same FSG), the only difference between grammatical and ungrammatical sentences was that the latter started with either TAM or JIC and ended with VOT or PEL, which were illegal endpoints. Infants were able to discriminate novel grammatical sentences from ungrammatical ones. In experiments 2 and 3, the authors tested whether infants were able to discriminate grammatical sentences generated by a FSG and ungrammatical sentences containing string-internal pairwise violations. Both grammatical and ungrammatical sentences began and ended with legal words but ungrammatical sentences contained illegal transitions according to the grammar (e.g., in the string VOT TAM PEL RUD JIC, VOT TAM, PEL RUD, and RUD JIC are illegal transitions). Also in this case infants successfully discriminated between grammatical and ungrammatical sentences. These results suggest that infants were able to extract the dependencies defined by the FSG. In the last experiment, the authors tested whether infants were able to abstract beyond specific word pairs. They trained infants on a FSG and then tested them on the same grammar using a novel lexicon. Infants were trained with sentences such as FIM SOG FIM
FIM TUP and tested on sentences such as PEL TAM PEL PEL JIC (grammatical) or *PEL RUD JIC VOT TAM (ungrammatical). Therefore, the grammatical structure was held constant across training and test phases while the vocabulary changed. The authors suggested that the observation that infants are able to discriminate between grammatical and ungrammatical sentences would indicate that they are able to abstract the grammatical structure beyond specific elements. Also in this experiment, they observed that infants were able to distinguish grammatical from ungrammatical sequences in the test phases. The authors concluded that infants showed significant abstraction capabilities.

Non-adjacent Dependencies

It has been claimed that non-adjacent dependencies are more difficult to learn than adjacent dependencies, since the learner has to keep track of the relation over irrelevant intervening material (Gomez, 2002). Therefore, the conditions facilitating the learning of non-adjacent syntactic regularities have been explored. Two conditions seem to favour the learning of this type of regularities, namely the context variability and the presence of perceptual cues.

The role of context variability has been investigated in a study performed by Gomez (2002). Participants were presented with three-word sentences (e.g., pel-kicey-jik) generated by one of two artificial languages (e.g., Language 1: aXd, bXe; Language 2: aXe, bXd). Since the adjacent dependencies between words were identical, participants had to track the relationship between the first and the third word to be able to distinguish the languages. X elements were drawn from different size pools (i.e., 2, 6, 12 or 24 elements). This difference aimed to test the influence of context variability (low variability or high variability) in the learning of non-adjacent dependencies. The author observed that learners, both adults and infants, were able to acquire the dependencies between non-adjacent words only when the intervening element was drawn from a big size pool, which created high context variability. This circumstance forced the participants to focus on the salient invariant structure in the sentence, that is the relation between the first and the third word. This result was replicated by Onnis and co-workers (Onnis, Monaghan, Christiansen, & Chater, 2004) showing that the best learning performances were obtain either at zero or at high variability.
Other studies investigating non-adjacent dependencies highlighted the fact that in order to learn this type of dependency, the presence of perceptual cues was necessary. In a study performed by Peña and colleagues (Peña et al., 2002), the authors investigated learners’ capacity to acquire and generalize a language called AXC-language in which the presence of a syllable A predicted C with a transitional probability of 1.0. In the training phase participants were presented with a continuous speech stream created by concatenating the trisyllabic pseudo-words generated by the AXC-language. Three “families” of AXC items were created. In each family the same AC pair was associated with three different intervening elements X (e.g., puliki, puRaki, pufoki). In the first experiment participants were tested on a word identification task in which they successfully chose the words of the language over part-words. In order to investigate whether participants only identified words in the stream (simple segmentation task) or whether they were also able to extract the structural regularities characterizing the stimuli, another experiment was performed. In experiment 2, participants were presented with the same type of training and tested in their capacity to identify rule words over part-words. Rule words were items that did not appear during the training but were consistent with the structure AXC. They were built up by using an intervening syllable X that was already present in the stream but not associated with those As and Cs (e.g., pubeki for the family puXki). The authors observed that participants failed to choose rule words over part-words, which suggested that they did not extract the structural regularities in the signal. However, as shown in experiment 3, if a subliminal pause (25 ms) was added after each AXC triplet in the training phase, participants successfully performed the task with rule words. Since the solely computation of transitional probabilities could not account for their results (as rule words did not appear during training), the authors suggested that the existence of another cognitive mechanism able to perform computations, probably of a rule-based nature, has to be acknowledged. The presence of perceptual cues, such as the pauses in these experiments, would allow the switch between one type of computational mechanism to the other.

Usually, non-adjacent dependencies in ALL studies focusing on syntax, mainly explored the learnability of embedded hierarchical structures generated by phrase structure grammars (e.g., Bahlmann et al., 2006, 2008; Fitch & Hauser, 2004; Friederici et al., 2006; Mueller et al., 2010; Perruchet & Rey, 2005; de Vries et al., 2011, 2008). Fitch and Hauser
(2004) performed an experiment in which they compared humans and non-humans capacity to learn a finite state grammar ((AB)\^n) and a phrase structure grammar (A^nB^n)^7. Their results highlighted that humans were able to discriminate both types of grammar, while cotton-top tamarins failed to learn the phrase structure grammar. The authors claimed that humans were able to process embedded hierarchical structures, whereas non-human primates were not. Subsequently Perruchet and Rey (2005) and de Vries and colleagues (2008) claimed that those results were not evidence of hierarchical structures processing but could be explained by alternative strategies such as counting. These results raised a debate about the learnability of embedded hierarchical structures in artificial grammar learning studies.

Interestingly, it has been recently proposed (Mueller et al., 2010) that embedded hierarchical structures (e.g., the rat the cat ate died) may be learned under specific circumstances, namely the presence of prosodic cues (i.e., pauses) marking the boundaries of main units. Mueller and colleagues investigated the learnability of a phrase structure grammar under different conditions and controlling for repetitions, in order to rule out the usage of counting strategies. The main difference between conditions was that the stream played during the training was characterized by the presence of a pause marking the main units (e.g., for the sequence dugikutodedetotu; condition 1: dugikuto_dedetotu; condition 2: du_giku_to_de_deto_tu, were the symbol “_” represents the pause). They compared these conditions with a control one (condition 3) in which the training was constituted by unsegmented speech. They showed that learning was possible only in the conditions characterized by segmented speech (conditions 1 and 2). The results of this study suggest that hierarchical dependencies are learnable when prosodic cues are included in the stimuli and that distributional cues alone are not sufficient for learning. The authors stated that this observation might account for previous negative findings (Perruchet & Rey, 2005; de Vries et al., 2008).

Recently, Saddy (2009) proposed the application of Lindenmayer Grammars to the study of human linguistics abilities. In particular, he exploited the learning of a Fibonacci

---

^7 A and B represent categories of CV syllables, finite state grammars generate sentences of the type ABABAB which are determined by local transitional probabilities, while phrase structure grammars generate sentences like AAABBBB, which cannot be described by local transitional probabilities.
Grammar: $A \rightarrow B \rightarrow AB$ which generates strings like: $A$, $B$, $AB$, $BAB$, $ABBAB$, etc. In his study, $A$ corresponded to the syllable /bl/ and $B$ corresponded to the syllable /ba/. Learners were presented with 3 minutes of unsegmented sequences of those syllables with flat intonation, following the Fibonacci grammar and subsequently tested in a discrimination task. Ungrammatical sequences were created starting from a random sequence not generated by the Fibonacci grammar. Four different symbol substitutions (e.g., $A$ is replaced by $AB$ and $B$ is replaced by $A$; $A$ is replaced by $ABA$ and $B$ is replaced by $AB$) were applied to this sequence in order to create four versions of ungrammatical items. It is important to notice that for the grammatical sequences, the transitional probabilities between adjacent syllables were close to random. In order to discriminate grammatical from ungrammatical sequences, participants had to abstract the structure-dependent regularities from the input. The fact that participants were indeed able to discriminate sequences generated by the Fibonacci Grammar from other sequences (percentage of correct responses varied from 60% to 68%), suggests that they were able to extract the hierarchical structure.

To sum up, in the previous sections we presented different types of structural regularities that have been investigated in artificial language learning experiments. Also in this case, as it was for segmentation, the learnability of both adjacent and non-adjacent dependencies has been explored. Interestingly, it has been shown that the presence of extra-cues (i.e., context variability, perceptual cues) was necessary for the learning of non-adjacent dependencies, while statistical regularities were sufficient to allow participants to learn adjacent dependencies. Finally, we highlighted a debate concerning the learnability of hierarchical structures in artificial languages. Crucially, our study develops within this final framework.

1.4.2.2. ALL in the non-linguistic domains

ALL paradigms have been used also with non-linguistic materials such as music tones, geometrical shapes, tactile stimuli, etc. The learning of adjacent and non-adjacent dependencies allowed researchers to highlight commonalities between linguistic and non-linguistic domains. Interestingly, as it has been observed for the linguistic domain, the learning of adjacent dependencies was easier than the learning of non-adjacent dependencies. The
learning of the latter required the presence of extra perceptual cues to be successful (Creel et al., 2004; Endress, 2010).

On the one hand, Saffran et al. (1999) showed that both infants and adults can readily calculate transitional probabilities between adjacent tones and use this information to group tones in word-like units in the same way as this has been shown in the linguistic domain (Saffran, Aslin, et al., 1996; Saffran, Newport, et al., 1996). On the other hand, Creel and co-workers (2004) observed that their participants were able to track non-adjacent dependencies only under certain circumstances. In their study, participants were exposed to a continuous stream constituted by concatenated triplets of tones. Each triplet of tones constituted a unit in which the transitional probability of adjacent tones was set to .5 and the transitional probability of non-adjacent tones was set to 1.0. In the test phase, participants were asked to determine which of two triplets (one was a correct triplet and the other was a misordered triplet) sounded more familiar. Participants performed the test at chance which suggested that they were not able to learn the relationship between non-adjacent tones. Subsequently, the authors tested whether the addition of a perceptual cue, namely the fact that non-adjacent tones had the same pitch, triggered the learning of the dependencies. Participants performed significantly above chance at the test showing that the presence of a perceptual grouping cue facilitated the learning of non-adjacent dependencies. These results illustrated that statistical learning is subject to constraints on the type of regularities that can be computed. As in the linguistic domain, non-adjacent dependencies were readily learned when they were determined over consonants or vowels, but not when they were determined over syllables (Newport & Aslin, 2004), so in the music domain, learning non-adjacent relations was possible only when perceptual grouping cues were available.

Non-linguistic materials have also been used to study whether attention played a role in the learning of non-adjacent dependencies. In a recent study, Pacton and Perruchet (2008) proposed an attention-based model to account for the learning of non-adjacent dependencies. In their experiments participants were faced with a series of digits sequences (e.g., 5 8 5 6 4 3 9 7 1 0 8 6 4 9 3 0 7 2 2 1) containing both adjacent (e.g., 4 follows 6) and non-adjacent regularities (e.g., 7 follows 3 after a variable number of intervening digits). Participants were required to read the sequence looking for a target digit (e.g., 9). When they detected the target,
a group of participants had to calculate the difference between the two digits preceding the target, and another group had to calculate the difference between the two digits surrounding the target. These different instructions were aimed to drive the attention either to the adjacent dependency or to the non-adjacent dependency. Subsequently participants were tested in a recognition task, they had to judge which one of two strings of three digits was part of the sequences previously presented. Half of the test stimuli were aimed to assess the learning of the adjacent dependency and half the non-adjacent dependency. The results showed that the type of dependency learned by participants depends on the direction of their attention during processing. Participants who were asked to pay attention to the preceding digits learned the adjacent dependency, but not the non-adjacent one, while participants who were asked to pay attention to the surrounding digits learned the non-adjacent dependency but not the adjacent one. These results support the idea that attention is a necessary condition for learning to occur, which has been proposed by models of learning (see Kruschke, 2003 for a review on this issue). Importantly, the authors suggested that this attention-based model could also account for previous ALL results, such as the ones from Creel et al. (2004). The observation that non-adjacent dependencies are learned when the materials possess specific properties (e.g., perceptual grouping cues) could be explained by the fact that the presence of these properties facilitates the attentional processing on the relevant events.

In sum, ALL in non-linguistic domains showed important similarities with ALL in the linguistic one. These observations suggest that, at least to a certain extent, there are general learning mechanisms that subserve knowledge acquisition in different domains. The fact that learners are equipped with a learning mechanism able to process both adjacent and non-adjacent dependencies in several cognitive domains leads to the question about how to constraint this powerful learning device. In the following section we will address this issue.

1.4.2.3. Learning Constraints

The existence of a cognitive mechanism able to track adjacent and non-adjacent dependencies between units of different nature would force learners to deal with a large number of computations. How do learners know which data are relevant for learning and which are not? How do learners know what kind of statistical information should be tracked?
The evidence that statistical learning plays a role in language acquisition has raised the problem of identifying which constraints may guide this powerful mechanism. The existence of constraints that should aid the learners in the acquisition of language is fairly widely accepted and there is general consensus about their innateness (Saffran, Aslin, & Newport, 1997; Yang, 2004). However, the nature of these constraints is still debated. In this section we will present the approaches that, in our opinion, summarize the debate about the nature of the constraints that allow learners to acquire language. We classify them as domain-specific and domain-general constraints.

**Domain-specific constraints:** To our knowledge there are no studies directly implementing Universal Grammar properties in order to investigate domain-specific constraints to higher level statistical computations in ALL. However, the role of other linguistic constraints on statistical computations was observed. In order to exemplify the role of domain-specific constraints in statistical learning, we have to compare a study performed with an artificial language (Saffran, Aslin, et al., 1996) and a study performed with natural language corpus (Yang, 2004). As we have mentioned in the previous section, it has been shown that adults and infants are able to segment a continuous speech stream based on the transitional probabilities between CV syllables (Aslin et al., 1998; Saffran, Newport, et al., 1996). Yang (2004) tried to replicate this result with a computational model tested on child-directed English, however he did not obtained the same results (i.e., the model was not able to efficiently extract the words from the corpus). How could the difference between the performance of the model and infants’ performance be explained? In order to be effective, the learning algorithm of the model needed to be constrained. The model’s performance improved if prior linguistic knowledge about phonological structures was added. The constraint applied in this study was that each word could have only one primary stress. Therefore, assuming the presence of this constraint, once the learner is presented with a sequence such as “bigbadwolf” (the bold font represents the stress), she automatically knows that it has to be segmented in three words. However, when the constraint fails to establish word boundaries, such as in the sequence “targetgramm”r the learner uses statistical learning to calculate local minima and determine the boundaries\(^8\). Since it has been shown that infants as young as 7 months are already

---

\(^8\) Word boundaries are assumed to be at “local minima”, that is the position in the stream were the TP is lower.
sensitive to the strong/weak patterns of prosody in natural language (Jusczyk, 1999), this type of domain-specific information represents a plausible candidate to constraint computations leading to word segmentation. In conclusion, Yang (2004) claimed that linguistic knowledge can help statistical learning in word segmentation providing domain-specific constraints to its application.

Two studies exploring word segmentation (Bonatti et al., 2005) and ABA pattern extraction (Toro et al., 2008) (see Section 1.4.2.1. for details about the experiments) showed the role of other linguistic constraints on ALL. The authors hypothesized that consonants and vowels in words carried different kinds of information. Therefore, depending on the type of task in which learners were involved (word identification vs. patterns generalization) the units over which the statistics were calculated varied. Transitional probabilities were calculated over consonants for the word identification task and over vowels to extract generalizations. The authors explained their results referring to the role of consonants and vowels in natural language. As proposed by Nespor and colleagues (2003), it is possible that consonants and vowels serve partially different roles in language processing. There are qualitative and quantitative differences between consonants and vowels (e.g., consonants are more than vowels, vowels are the carrier of grammatical intonation, the brain localizations for consonants and vowels differ, etc.) which suggest that consonants may contribute to the identification of the lexical elements, while vowels may cue regularities and contribute to the interpretation of grammar. All in all, these studies showed that linguistics information works as (domain-specific) filter on the statistical computations the learners perform in order to acquire the lexicon and aspects of syntax.

Domain-general constraints: It has been proposed that constraints on language learning may arise from general characteristics of the human cognitive architecture, memory, perception and prior experience (Saffran, 2002). For instance, Mehler and collaborators suggested that humans (and maybe non-human animals) are equipped with perceptual or memory primitives (POMPs) which constrain statistical learning (Endress & Mehler, 2010; Endress, Dehaene-Lambertz, & Mehler, 2007; Endress, Nespor, & Mehler, 2009; Mehler, Nespor, Endress, & Gervain, 2008). These mechanisms, originally non-linguistic, are recruited by the language
faculty and act as filters which constraint the types of regularities that are learnable. In this way they simplify the input for other learning mechanisms such as statistical learning.

A POMP which has been shown to be relevant for language learning is the one processing identity relations (i.e., repetitions). Repetition (or reduplication) seems to be a useful feature for computing complex sequences such as those that exist in language. Child-directed speech, in fact, is characterized by repetitions of words and utterances (see Snow, 1972 for a characterization of child-directed speech), which make this primitive a good candidate as learning constraint. A repetition-based mechanism can also explain results to ALL experiments such as the one proposed by Marcus et al. (1999) (see Section 1.4.2.5. for details about the experiment). Since the grammars used in their study entail repetitions (i.e., ABA, ABB, AAB), their detection and the fact that they were at the edges of the sequences might have triggered the learning of the grammars. For instance, Endress, Dehaene-Lambertz and Mehler (2007) observed that only when structure regularities were characterized by repetitions, adults were able to learn them. In their study they compared the learning of two simple grammars using triplets of piano tones. The first grammar was based on the repetition of the elements in a sequence (e.g., ABA vs. ABB) and the other was based on “ordinal” relations between the pitch of the elements (e.g., Low-High-Middle vs. Middle-High-Low). Participants showed a good performance in the task (i.e., determine whether two sequences corresponded to the same grammar) only for the repetition-based grammar (79% correct responses) and not for the “ordinal” grammar (53% correct responses). Therefore, a perceptual based mechanism seems to guide (constraint) the learning of simple grammars.

Several other studies pointed out that learning might also be constrained by the presence of overlapping cues of different nature (e.g., statistical, prosodic, phonological). Therefore, the integration of these multiple cues might guide language learning in a more reliable way than the presence of only one type of cue (e.g., Monaghan & Christiansen, 2008). Indeed, corpora studies (e.g., Kelly, 1996; Monaghan, Christiansen, & Chater, 2007; Shi, Morgan, & Allopenna, 1998) and ALL experiments (e.g., Monaghan, Chater, & Christiansen, 2005; Newport & Aslin, 2004; Peña et al., 2002) showed that learners benefit from the presence of multiple cues in speech, for segmentation as well as for structure regularities computations.
For example, Christiansen and co-workers (Christiansen, Allen, & Seidenberg, 1998; Christiansen, Conway, & Curtin, 2005) proposed a connectionist model for word segmentation which integrated different types of cues, such as sequential phonological regularities, lexical stress, prosodic boundaries. They investigated the performance of simple recurrent network both with artificial and natural language input (e.g., Allen & Christiansen, 1996). In one of their studies (Christiansen et al., 1998), they trained a simple recurrent network with a corpus of child directed speech. The authors compared the performance to a segmentation test of networks trained with information about phonology, utterance boundaries, and lexical stress with the performance of networks trained only with information about phonology and utterance boundaries. Their simulations showed that the network provided with three cues was able to segment 23 out of 50 words while the network provided with two cues was able to segment 11 out of 50 words. These results supported the idea that the presence of multiple cues guides segmentation in a more reliable way.

To sum up, the presence of (innate) constraints to the mechanism tracking regularities involved in language acquisition is widely accepted. Nevertheless, the nature (domain-specific vs. domain-general) of these constraints is still matter of debate.

1.4.2.4. Generalization

ALL studies have been used to explore different types of generalization and the implications that this process has in determining the nature of knowledge resulting from learning. In particular, while certain types of generalization have been extensively studied with ALL paradigms, others have been investigated only by a very few studies (i.e., generalization to novel syntactic structures). We will organize this section around four types of generalization investigated in ALL studies: the pattern-based generalization, the category-based generalization, the generalization to novel syntactic structures and the across modalities generalization.

The first type of generalization investigated with ALL studies is called pattern-based generalization. As Folia (2010, p. 193) stated, pattern-based generalization “can be described in terms of relations over surface-based (e.g., physical) characteristics of the stimuli”. Therefore, this type of generalization is based on learners’ capacity to detect relationships
(e.g., same/different) between the elements constituting the sequences of their acquisition set (e.g., in a sequence generated by the pattern ABA learners have to detect that within the sequence the first and the last elements are identical). Evidence of this kind of generalization came from studies using linguistic and non-linguistic stimuli both with adults and infants learners (e.g., Gomez & Gerken, 1999, 2000; Marcus et al., 1999; Saffran, Pollak, Seibel, & Shkolnik, 2007). For instance, Gomez and Gerken (1999) explored 1-year-old infants’ capacity to perform pattern-based generalizations (see Section 1.4.2.1.2. for details about the study). In their study, infants were able to discriminate between grammatical sentences generated by a FSG and ungrammatical sentences even when the lexicon between training and test phases was different. The authors proposed that infants might have based their judgments on perceptual characteristics of the materials. It has been claimed that since finite state grammars allow paths to create loops, learners could have detected characteristic repetitions patterns in the input. For example, if we compare the sentence PEL TAM PEL PEL JIC with the sentence FIM SOG FIM FIM TUP, we can identify a repetition pattern such as x_xx_ in which “x” represents the same word. The extraction of these patterns may therefore explain the good performance when a novel lexicon is used in the test phases since learners can identify in the test strings the repetition patterns they were exposed to during the training. Interestingly, researchers have observed that when these repetition patterns were not available, because the finite state grammar did not contain any loop, the generalization was not possible (e.g., Gomez, Gerken, & Schvaneveldt, 2000; Tunney & Altmann, 2001). Therefore, the theoretical approach supporting the study of this type of generalization reduces the knowledge transfer to a process merely based on the detection of the repetition patterns that are similar between the input in the familiarization and that constituting the test.

*Category-based generalization* is described in terms of operations over abstract variables. It determines the generalization of some kind of regularities abstracted from an initial set of items to new items belonging to the same category. In order to understand how this type of generalization differs from pattern-based generalization let us consider the following example. We compare the pattern-based representation ABA and the category-based representation Noun-Verb-Noun. Although in both cases we have to refer to the identity relation in order to recognize the regularities, in the former case abstracting the pattern from a stimulus such as
wofewo requires noting that the first element is physically identical to the third element. In the latter case the identity relation is determined over the abstract categories, namely the learner has to notice that both the first and the third elements belong to the same category “Noun”. It has been claimed that this type of generalization is useful in language learning since once the learners have recognized that a new word belongs to a certain category, all the syntactic properties associated to the other words of the same category are automatically applied to the new one. Given the importance of this type of generalization for language learning, several studies investigated how learners determine that a word belong to a certain category (e.g., Braine et al., 1990; P. J. Brooks, Braine, Catalano, Brody, & Sudhalter, 1993; Frigo & McDonald, 1998).

Items categories can be induced on the basis of their position within a sequence (e.g., all items following a in a sequence like aX will be considered as belonging to the X category, whereas items following b in sequence like bY will be considered as belonging to the Y category). However, since determining the membership to an abstract category is a difficult task, it has been argued that, if the elements belonging to the same category are conceptually or perceptually marked, the learning task is easier. For example, Frigo and McDonald (1998) showed that learning was facilitated when category membership was phonologically marked (e.g., all members of category I ended with –ash and all members of category II ended with –org) than when it was not marked.

In a similar study, Gerken, Wilson and Lewis (2005) tested infants’ capacity to form syntactic-like categories (i.e., feminine, masculine) in an ALL study. 18-months-old American infants were trained with a miniature version of Russian gender paradigm in which feminine words ended with –oj and –u (e.g., polkoj, polku), and masculine words ended with –ya and –yem (e.g., zhitel’ya, zhitel’yem). During the training phase some words were presented only in one of the two possible versions of their gender (e.g., only ruchkoi was presented and not ruchku). This allowed the experimenters to use the remaining version (e.g., ruchku) as novel

---

9 It is important to notice that for the category-based generalization, the perceptual similarity (or identity) is between members of the same category (e.g., all members of category I end in –ash: oprash, dimash, yerash) and not between elements in the same sequence (e.g., as in VOT PEL PEL JIC) as it is for the pattern-based generalization.
grammatical test items\textsuperscript{10}. Ungrammatical test items were built up using the incorrect gender marking on the roots previously heard by infants (e.g., ruchkyem). Results to the head-turn preference task showed that infants were able to discriminate new grammatical items from ungrammatical ones. The performance of a control group that did not perform the training was also measured. These infants did not show discrimination between grammatical and ungrammatical items. Therefore the authors claimed that the discrimination effect observed for the experimental group was indeed due to the training with the gender paradigm. Infants were able to differentiate elements belonging to two separate syntactic-like categories generalizing the knowledge about the way gender was marked that they acquired during the training.

To our knowledge very few studies in ALL investigated learners’ ability to generalize their knowledge about syntactic regularities to novel syntactic structures which can be induced on the basis of the initial structures on which they have been trained. Lany and co-workers (Lany & Gomez, 2004; Lany, Gomez, & Gerken, 2007) performed two studies in which they investigated this type of generalization. In their studies adult participants were trained on \textit{aX}, \textit{bY} artificial languages. In order to learn these languages participants had to learn to categorize pseudo-words in categories \textit{a}, \textit{b}, \textit{X}, and \textit{Y} (see category-based generalization) and to learn the distribution of the elements within the sequences, namely that an element \textit{a} (e.g., ush, dak) always predicted an element \textit{X} (e.g., lepit, zamit, etc.), and an element \textit{b} (e.g., ong, rud) predicted an element \textit{Y} (e.g., bivul, jerul, etc.). Subsequently, participants trained with \textit{aX} and \textit{bY} sequences (e.g., ush lepit, ush zamit, rud bivul, rud jerul) were tested in a grammaticality judgment task on sequences of the type \textit{acX} and \textit{bcY} (e.g., ush hes lepit, ush kaf zamit, rud hes bivul, rud kaf jerul). The performance of a control group that did not perform the training with \textit{aX} and \textit{bY} sequences was also measured. The authors observed that participants who performed the training were able to generalize their knowledge about the co-occurrence of elements belonging to the category \textit{a} with elements belonging to the category \textit{X} and the co-occurrence of elements belonging to the category \textit{b} with elements belonging to the category \textit{Y} also when an intervening element (i.e., \textit{c}) was added between them (mean d-prime score between block 1 and block 2 = 1.25). On the contrary, participants that did not perform the

\textsuperscript{10} In this context the terms “grammatical” and “ungrammatical” refer respectively to the fact that the gender was correctly marked or to the fact that the gender was incorrectly marked.
training were not able to learn the \( acX, bcY \) languages (mean d-prime score between block 1 and block 2 = 0.24). Therefore, participants were able to learn to associate an element \( a \) adjacent to an element of the category \( X \). They were also able to generalize the knowledge of this dependency to a non-adjacent dependency in which the element \( a \) and the element belonging to the category \( X \) were separated by an intervening element \( c \). In the present dissertation we developed the idea of generalization to novel syntactic structures employing different types of category marking and using more complex structure dependencies than the one employed in the above-mentioned experiments.

The *across modalities generalization* has been investigated in ALL paradigms, however the existence of such type of generalization is still a matter of debate. Generalization across modality occurs when regularities trained in a certain modality/domain (e.g., visual) are tested in a different one (e.g., auditory). Early experiments (Altmann, Dienes, & Goode, 1995; Dienes & Altmann, 1997) showed that transfer between domains was possible. In these experiments participants were usually trained with simple finite state grammars and then asked to classify the test sequences on whether they were created according to the same grammar or not. Altman, Dienes and Goode (1995) performed a series of experiments in which they trained adults with sequences in a certain modality (e.g., tones, spoken syllables, graphic symbols) and subsequently tested them with sequences belonging to a different modality. For example, in the first experiment participants were presented with sequences of tones generated by a finite state grammar. In the test phase they performed a classification task in which they had to decide whether sequences of letters presented in a different modality (i.e., visual modality) were grammatical, that is they were generated by the same finite state grammar as the training items, or not. The performance of these participants was compared to the performance of a control group that did not perform training. Results showed that participants who previously performed the training phase had a better performance in the test compared to the control group. According to the authors, participants were able to correctly classify the test sequences, even if their modality was different compared to the modality of those in the training phases, relying on a modality independent representation of the knowledge acquired during training.
The approach supporting the existence of across modalities generalization received novel evidence from recent studies (Hupp & Sloutsky, 2011; Hupp, Sloutsky, & Culicover, 2009). In an elaborated study, Hupp and Sloutsky (2011) explored across modality generalization in 8- and 16-months-old. In the first experiment it was established that infants had a preference to detect changes at the beginning of a visual sequence. In fact, when familiarized with a sequence of two shapes of the same colour (e.g., familiarization: green triangle – green cube) and then tested with sequences whether an element was added to either the beginning (e.g., condition Pre+: orange musical note – green triangle – green cube) or at the end of the familiarization sequence (e.g., condition Post+: green triangle – green cube – orange musical note), participants looked longer test sequences in the condition Pre+. The same experiment was conducted in the auditory modality (i.e., using spoken syllables sequences, familiarization: Ki-Tu, Pre+: FO-Ki-Tu, Post+: Ki-Tu-FO) showing that infants did not have any preference for the beginning or the end of the sequence and that they were equally likely to notice changes in both conditions. In the last experiment the authors trained infant to pay attention to the end of the visual sequences. In order to do so, the second element (e.g., blue heart) of the training visual sequence (e.g., blue sun – blue heart) was made salient by doing something like jumping, tilting, expanding, etc. After training, infants were familiarized to the auditory sequence (e.g., Ki-Tu) and then tested in the Pre+ and Post+ conditions with auditory materials. Results showed that 16-months-old, but not 8-months-old, looked longer to the Post+ sequences, suggesting that they were able to transfer the knowledge acquired in the visual modality (i.e., focus on the end of the sequences) to the auditory one. Therefore, these results added evidence to the approach supporting the existence of across modalities generalization. The authors accounted for the fact that only older infants were able to transfer their knowledge from one modality to the other, suggesting that transfer broader with development as children become more experienced learners.

A contrasting approach suggests that across modalities generalization is not possible as the representations resulting from learning are modality-dependent (i.e., stimulus-specific) and not abstract (e.g., Bly, Carrión, & Rasch, 2009; Conway & Christiansen, 2006; Johansson, 2009). For example, Conway and Christiansen (2006) investigated whether learners encode sequential patterns in abstract or stimulus-specific representations. They exposed participants
to visual sequences of colours generated from a finite state grammar (Grammar A) and to auditory sequences of tones generated from a different finite state grammar (Grammar B). Subsequently, in the test phase, participants were asked to perform a classification task. For one group of participants all test items for both grammars were visually presented (i.e., colours were used) and for the other group all test items for both grammars were auditorily presented (i.e., tones were used). The authors hypothesized that, if participants’ knowledge was stimulus-specific, they should have been able to correctly classify sequences only if they were presented in the same modality as the training sentences generated from the same grammar (e.g., participants tested in the visual modality should be able to classify sequences generated by the grammar A and not sequences generated by the grammar B). If participants’ knowledge was abstract they should have been able to classify all the sequences generated by both grammars disregarding the presentation modality. As their results followed the former pattern, the authors concluded that participants’ learning resulted in a stimulus-specific knowledge and no transfer across modalities was possible. Clearly, these results challenged previous findings suggesting that artificial grammar learning resulted in modality-independent (Altmann et al., 1995) or abstract representations (Reber, 1989). Conway and Christiansen explained this fact by pointing out that those transfer data did not necessarily imply an abstract representation of the knowledge. Participants could have successfully performed the task by simply encoding the repetition patterns characterizing the finite state grammar.

1.4.2.5. Types of Knowledge

The concept of generalization is tightly linked to the nature of the output that results from learning. This link is due to the fact that generalization may result from knowledge of different types. Different theories have been proposed to account for the nature (format) of the knowledge generated from ALL. In the present section we will introduce four theories, i.e., abstract knowledge, exemplar, fragment, and perceptual primitives theories, which are in our judgment the most representative (see Pothos, 2007 for a comprehensive review). These theories have been mainly developed to describe the type of knowledge resulting from artificial grammar learning studies (in particular those involving the learning of finite state grammars). Nevertheless, they are a valuable support to account for the knowledge generated by other kinds of artificial grammars.
Rule theories – The idea common to these theories is that learning in ALL studies implies the construction of abstract mental representations\(^\text{11}\). These representations are supposed to describe the grammar generating the training sequences and to account for the legal (grammatical) sequencing of the elements (e.g., words) in those strings (Shanks, Johnstone, Kinder, French, & Cleeremans, 2002). The definition of rule knowledge is not univocal. According to the theories describing the knowledge resulting from artificial grammars learning experiments, such as those employing finite state grammars, the definition of rule refers to grammaticality. Participants are supposed to extract the knowledge in the form of a network of rules that refers to the deep structural characteristics of the input (Reber, 1989). Evidence in favour of this theory mainly came from the studies showing generalization. In these studies, training and test items were characterized by a different lexicon (e.g., set of letters such as in Gomez and Gerken (1999)) or different modalities (e.g., training with letters and test with tones, as in Altman et al. (1995)). Researchers suggested that learners had to perform their judgment on the basis of abstract knowledge, as no other common perceptual features between training and test materials were available.

According to Marcus and co-workers (1999), the knowledge resulting from ALL is stored in the format of algebra-like rules. These rules are defined as open-ended abstract relationships between variables for which we can substitute arbitrary items. In their study, 7 month old infants were presented with three-word sentences generated by the pattern ABA (e.g., ga ti ga) or ABB (e.g., ga ti ti). Subsequently, they were tested with sentences consistent or inconsistent with the pattern they had learned, but built up by new words (e.g., wo fi wo). Infants were able to discriminate consistent from inconsistent sentences. In order to rule out the hypothesis that this discrimination was based only on a surface cue like reduplication, the authors run another experiment comparing ABB sentences vs. AAB sentences. The same result, namely discrimination was obtained. Since the lexicon used in the test phases was different from the one used during training, infants could not rely on the computation of transitional probabilities (transitional probabilities were set to 0) to discriminate the sentences.

\(^{11}\) Two meanings of “abstract” are appropriate in the context of this account. Johansson (2009) has defined abstract representations as surface-independent representations which are amodal and can be disentangled from the surface properties of the items. Redington and Chater (1996) have offered another description of abstract representations defining them as the knowledge of the deep structure characterizing the input.
The authors stated that only assuming that the output of the learning process had the form of an abstract algebra-like rule was it possible to account for these results.

Several criticisms have been addressed to the interpretation of this study and a debate arose on whether infants are able or not to learn something that goes beyond the statistical structure of the training items. In particular, Seidenberg and Elman (1999) stated that “the conclusion by Marcus et al. that the infants had learned rules rather than merely statistical regularities is unwarranted”. They proposed that instead of calculating statistics concerning the co-occurrence of certain elements (e.g., wo + wo + fe), learners might have calculated statistics over relationships of the type “same-different” (e.g., such as the first element is followed by an identical one and then by a different one). Therefore, Marcus and colleagues’ study results could simply indicate that learners were able to compute different types of statistics rather than indicate that learners were able to acquire abstract rules.

McClelland and Plaut (1999) developed this concept proposing a broader notion of statistics which was not limited to the computations of transitional probabilities between particular elements. The idea of statistics that emerged from their response to Marcus et al. was, as Marcus (1999) pointed out, “that any kind of relation between any kind of information, concrete or abstract” could be statistically tracked. Statistical learning mechanisms were therefore sufficient to explain the data. However, as Marcus outlined in his response to this paper (Marcus, 1999), such a broad definition of statistics includes the idea of rule. For example, if in a language all sentences are constituted by a noun-phrase followed by a verb-phrase, the language can be described in terms of phrase-structure rule as follows: Sentence → Noun phrase, Verb phrase. It could also be described in terms of statistical patterns as the verb phrase follows the noun phrase with a transitional probability of 1.0. Marcus and co-workers stated that the definition of statistics they were arguing against only referred to the one limited to the calculation of transitional probabilities between words, and not to this broader definition. We can conclude that the difficulty in settling this debate resides in the fact that an ambiguous definition of the levels of abstraction over which statistics can be computed has been adopted in the literature.


Chapter 1. Theoretical Framework

Exemplar theories – According to this account, learners’ knowledge resulting from artificial grammar learning, consists of relatively unprocessed representations of whole strings. This representation allows the learner to encode the (perceptual) similarity between test and training items. During the test phase, learners are supposed to retrieve from their memory representations of specific training items and compare the test items with those examples. The grammaticality of a test string is then calculated on the basis of its similarity to the training items. Similarity can be defined as “the mental operation that allows us to compare two instances and establish their commonalities” (Pothos, 2007, p. 228). For instance, Brooks and Vokey (1991) defined similarity as the number of letter changes between two strings. A test item was “similar” to a training item if they differed only by one letter while it was “dissimilar” if it differed in more than one letter from the training item. In an artificial grammar learning study, Vokey and Brooks (1992) investigated the salience of similarity in a transfer task (i.e., participants were tested with novel grammatical and ungrammatical items generated from the same finite state grammar). Grammatical and ungrammatical items were also classified along the similarity dimension. For example, if the training item was MXR, the similar grammatical test item was MVR while the similar ungrammatical test item was MTR; dissimilar grammatical and ungrammatical test items were also created substituting more than two letters from the corresponding training items. Results showed that items were more likely judged as grammatical if they were grammatical or similar to the training ones. Therefore, according to this model learning involves the computation of similarity between novel and old presented items.

Fragment theories – An alternative theory based on the representation of stimuli surface features has been proposed by Perruchet and Pacteau (1990). The authors claimed that knowledge resulting from artificial grammar learning is based on the memorization of fragments of the training items without any rule abstraction. Test items are judged as grammatical if they contain the fragments presented during training. A fragment is defined as a part of a training item. Since studies supporting this theory were conducted using strings of letters as training items, fragments were usually bigrams or trigrams. In an artificial grammar

---

12 The main effect of grammaticality was discussed by the authors referring to the fact that a novel grammatical item was moderately similar with many training items. For instance, the new grammatical item MVR is similar to MXR and also moderately similar to MVRVM, VMRMVR.
learning paradigm (Perruchet & Pacteau, 1990), participants were asked to perform a grammaticality judgment task. They were divided into two groups. Participants belonging to the first group received a standard training phase constituted by a list of strings of letters generated by a finite state grammar. Participants belonging to the second group were only presented with bigrams (e.g., VR, RX, VT, etc.). These bigrams were obtained by segmenting into pairs of letters the strings generated by the finite state grammar (and presented to the first group). Subsequently, both groups performed the grammaticality judgement task on the same set of test strings (grammatical test strings were novel compared to the training strings but generated by the same finite state grammar with the same set of letters; ungrammatical strings were constituted by the same set of letters in positions that were not permissible according to the grammar). Results showed that the performance of the second group (61% correct responses) was equivalent to the performance of the first group (59% correct responses). Participants exposed only to bigrams performed equally to participants exposed to the whole strings. According to the authors this result suggested that bigram knowledge was sufficient to account for the performance in typical grammaticality judgment tasks used in artificial grammar learning studies. Therefore, according to this account, the knowledge resulting from artificial grammar learning is based on the knowledge of the fragments constituting training items, rather than on the knowledge of a complex system of abstract rules as it was suggested by Reber (1989). As it has been pointed out (L. R. Brooks & Vokey, 1991; Pothos, 2007), this theory is very similar to the interpretation of artificial grammar learning offered by Dulany et al. (1984). According to these authors, artificial grammar knowledge is stored in the form of “micro-rules” which are nothing but rules describing parts of the sequences.

Perceptual primitives account – This approach tries to reconcile the theories concerning the form of the knowledge that we presented above. As we presented them (and as they are usually presented in the literature), learners either extract statistical regularities or rule-based regularities. These processes are usually considered as incompatible. The approach proposed by Endress (2005) aims to synthesize the statistical nature of the learning output and the existence of abstract rules. In his dissertation, Endress suggested that the generalization observed in Marcus et al. study (1999) might be performed by symbolic operations specialized in the processing of repetitions (e.g., ABB) and edges (i.e., the fact that repetitions are at the
edge of the sequence). These symbolic operations are called *perceptual primitives* and are defined as simple operations that work under specific circumstances (i.e., repetitions and edges). According to this account, perceptual primitives are proposed to be the learning constraints that applied to statistical operations allow the learning of symbolic representations. In a series of experiments, Endress, Scholl and Mehler (2005), focused on the role of the perceptual primitives as “processing that could underlie the extraction of symbolic structure” (Endress et al., 2005, p. 408). They compared the learning of regularities of the form ABCDEF and of the form ABCDDEF. The difference was determined by the position of the repetition. In the former case, since it was at the edge of the sequences, it was more salient than in the latter. Results showed that learners performed successfully to the generalization task (i.e., they were tested on the same structure but with a novel set of syllables) only when the repetitions were located at the final edge of the sequences. According to the authors these results support the idea that generalization is due to constrained symbolic operations (i.e., perceptual primitives) that are triggered by the perceptual salience of certain type of structures (i.e., repetitions and edges). This view is therefore to be preferred to the view proposing the existence a general formal process that computes all possible regularities equally well in the manner of a computer, as suggested by Marcus (2001)\(^\text{13}\).

To sum up, different theories have tried to account for the nature of the knowledge resulting from ALL. These theories can be traced back to four main accounts. The first assumes that the learning results in abstract mental representations (i.e., rules); according to the second account, learning results in the storage of training examples and the computation of the perceptual similarity between training and test items; according to the third theory, learning results in the memorization of parts (i.e., fragments) of the training items. Finally, the last approach outlined in this section proposes a synthesis between statistical and symbolic operations by the means of perceptual primitives. The debate about the format of the knowledge is still ongoing and novel evidence is necessary to clarify which account is the most adequate.

\(^{13}\)“Variables are one part of the story, operations over those variables another. To clarify the difference, consider the distinction in digital computers between registers and *instructions*. Registers store values; instructions, such as “copy” and “compare,” manipulate those values. My hunch is that the brain contains a similar stock of basic instructions, each defined to operate over all possible values of registers.” (Marcus, 2001, p. 58)
1.4.2.6. Summary

Artificial language learning studies have been extensively used to evaluate the role of different learning mechanisms that may be exploited in language acquisition and processing. The general idea at the basis of these experiments is to investigate the acquisition of several types of regularities in a simplified and controlled environment. In the previous sections we first presented the main methodological aspects of artificial language learning experiments with particular attention to the types of languages, the types of experimental tasks and the population tested in these studies. These sections allowed us to introduce the tools the reader may use to better comprehend the literature review subsequently presented. The review of the topics explored in ALL allowed us to introduce some aspects of the artificial language we will use in the experiments constituting this thesis. As it will be outlined in detail in the next section about research questions we aim to contribute to some open debates. In particular, we will investigate whether human adults are able to learn rules based on the hierarchical organization of the elements, which cues allow this learning and which types of knowledge results from the learning.
1.5. Research Questions and Empirical Framework

1.5.1. Research Questions

The artificial language learning paradigm has been used to study the cognitive processes involved in language acquisition. These artificial languages are based on rules mimicking aspects of natural language structures. Some studies focused primarily on the mechanisms used by infants and adults to perform lexical segmentation; other studies focused on syntactic relationships learning. Among the latter, several compared grammars inspired by Chomsky’s hierarchy (e.g., Finite State Grammars vs. Phrase Structure Grammars) (see for example Bahlmann et al., 2006, 2008; Fitch & Hauser, 2004; de Vries et al., 2008 among others). The learning of rules based on ABA – ABB structures has been also explored (Gervain, Macagno, Cogoi, Peña, & Mehler, 2008; Marcus et al., 1999). More recently, artificial languages have been employed to study long-distance dependencies and to explore regularities extraction between non-adjacent elements (see for example Creel et al., 2004; Gomez, 2002; Newport & Aslin, 2004; Newport et al., 2004; Peña et al., 2002 among others). The learnability of adjacent, linear based grammars is well established. On the contrary, despite the amount of data collected, evidence for the learnability of hierarchical syntactic structures in artificial languages is far from being conclusive.

The primary aim of this work was to contribute to the study of syntactic dependency learning. Two main purposes characterized the present work: the first one was to collect evidence about the learnability of different types of rules based on linear and hierarchical syntactic dependencies in an artificial language; the second purpose was to investigate the factors influencing grammatical rule acquisition. These questions can be formulated as follows:

Q1. Are adults able to learn syntactic dependencies in an artificial language
   a) When they are characterized as linear dependencies?

The learning of dependencies based on the position of the elements in the linear sequence of the items (or tone-words or geometrical shapes sequences, etc.) has been studied in ALL experiments. Two types of linear dependencies have been investigated: adjacent and
non-adjacent dependencies. On one hand, it has been shown that infants, adults and non-human primates easily track adjacent dependencies between items of different nature. On the other hand it seems that learning non-adjacent dependencies is more complex and requires specific cues (e.g., phonological cues). Moreover, to our knowledge none of the previous experiments investigated the learning of a rule generating both adjacent and non-adjacent dependencies. In the first two experiments of this dissertation we explored the capacity of adults to learn syntactic dependencies (adjacent and non-adjacent), generated by the linear subject-verb agreement rule in linguistic and music domains.

b) When they are characterized as hierarchical dependencies?

Previous artificial language studies have shown that adults are able to learn syntactic rules based on Finite State Grammars; however it is still not clear if they are also able to learn hierarchical structures (Phrase Structure Grammars) (Fitch & Hauser, 2004; Perruchet & Rey, 2005; de Vries et al., 2011, 2008). We created a hierarchical rule for subject-verb number agreement as a tool for the investigation of complex structure learning. In this work, we used a complex artificial language to study the acquisition of structure-dependencies. Our purpose was to investigate whether human adults are able to extract and generalize syntactic regularities generated by a hierarchical rule and which factors influence this learning.

Q2. What are the factors that influence rule learning in an artificial language?

Three main factors have been taken into account in this dissertation and their influence in the learning of rules implemented in artificial languages was investigated in six experiments. First, we studied whether the type of units used in the artificial language (linguistic vs. non-linguistic) influences the learning of the agreement rule. Second, we investigated which phonological and prosodic cues in the signal influence the creation of relationships between the elements involved in the syntactic rule. This investigation allowed us to explore whether a complex syntactic rule requires several co-occurring cues to be learned. Finally, we explored the influence of semantic cues in the learning of an artificial language. In the following paragraphs these factors are explained.
a) The types of units

In the review of the artificial languages (Section 1.4.1.1.) we already outlined the fact that different types of units are used in these experiments. In the present work we will focus on a comparison between units from different cognitive domains, namely music and language. Several studies (e.g., Jentschke, Koelsch, & Friederici, 2005; Koelsch, 2005, 2006; Koelsch, Gunter, Wittfoth, & Sammler, 2005; Patel, 2003) claimed that there are similarities between the processing in the linguistic and music domains. As a matter of fact, both music and language involve the organization of elements in structured sequences. A growing amount of evidence from neuroimaging studies have pointed out some kind of cerebral overlap between the systems underlying these two domains (Jentschke et al., 2005; Koelsch et al., 2002, 2005) suggesting the existence of cognitive processes shared between these domains. On the contrary, results from neuropsychology showed a double dissociation between musical processing and sentence processing (Dalla Bella & Peretz, 1999; Patel & Peretz, 1997). Therefore, the debate about whether language and music processing share cognitive mechanisms is still ongoing.

As previously mentioned (see Section 1.4.2.4.), artificial language learning paradigms have been exploited to investigate whether common general-domain processes are employed in these domains. To do so, the materials in these studies were either linguistic or musical items. In particular, it has been shown that transitional probabilities between adjacent elements can be easily computed between syllables (Saffran, Aslin, et al., 1996; Saffran, Newport, et al., 1996) as well as between tones (Saffran et al., 1999). However, the learning of non-adjacent dependencies in both the linguistic (Bonatti et al., 2005; Newport & Aslin, 2004; Toro et al., 2008) and the music domains (Creel et al., 2004; Endress, 2010) seems to be constrained by other properties of the materials. Experiments 1 and 2 aimed to further investigate the similarities and the differences between linguistic and music domains. The objective of these two experiments was to examine differences and similarities between the learning of regularities in language and music, and to study cross-domains generalization by the means of an artificial language. In order to do so, we compared the learning of the same linear subject-verb agreement rule with non-linguistic (i.e., tones) and linguistic (i.e., pseudo-words) materials (Experiments 1 and 2).
b) Phonological and prosodic cues

It has been shown that phonological cues represent an important source of information for learners during language acquisition, especially in grammatical category assignment (Durieux & Gillis, 2001; Frigo & McDonald, 1998; Kelly, 1996). In the present studies we created an artificial language in which grammatical categories and number were cued by different phonological cues (e.g., final vowels, place of articulation for consonants, number of syllables, syllables’ structures). We investigated the influence and the salience of these markers on the learning of a linear and hierarchical subject-verb agreement rule in an artificial language learning paradigm.

Since the hierarchical rule investigated here was based on the constituents’ structure of the sentences generated by the artificial language, it was necessary to use a cue marking this organization. We introduced also a prosodic cue (i.e., a pause) in order to mark constituents’ boundaries. Indeed, it has been shown that the presence of prosodic cues marking how words are grouped together in the sentence structure is fundamental for a successful acquisition of syntax (Morgan & Demuth, 1996a; Morgan & Newport, 1981; Morgan, Meier, & Newport, 1987). In our study, we explored the role of phonological and prosodic cues in the learning of the hierarchical subject-verb agreement rule (Experiments 3 to 6).

c) Semantic cues

It has been proposed that syntax and semantics interact in language acquisition. For instance, the semantic bootstrapping hypothesis has been proposed as a possible explanation of how infants are able to assign grammatical categories and the syntactic structure to the input they receive (e.g., Pinker, 1984; Wexler & Culicover, 1980). The major claim of the semantic bootstrapping hypothesis is that there is a close relationship between semantic and syntactic categories that can be exploited by infants during language acquisition. Semantic information may represent a critical source of information to distinguish categories; for example, nouns usually refer to objects, whereas verbs tend to refer to actions. Interestingly, it has been observed that in infant-directed speech there is good correspondence between semantic and syntactic information. For example, nouns referred to things or persons, verbs
referred to actions, spatial preposition referred to spatial information, etc. (Rondal & Cession, 1990). Semantic information may thus play an important role in language acquisition.

However, little research has been done to directly explore whether the integration of semantic cues would lead to a better learning of a language. To our knowledge only a few studies (Moeser & Bregman, 1972; Morgan & Newport, 1981; Morgan et al., 1987; Mori & Moeser, 1983) used artificial language learning paradigms to analyze the role of semantics in syntactic rule learning and they reached contrasting results. The majority of studies introducing semantics in the artificial languages employed did not directly investigate its role. Instead they implicitly assumed a facilitatory effect of semantic cues on learning. In our experiments we directly investigated the contribution of semantic cues in the learning and generalization of the hierarchical subject-verb agreement rule in an artificial language learning paradigm (Experiment 5 and 6).

In sum, in this section we presented the two questions we aimed to answer with the present dissertation. The first question concerns the capacity of human adults to learn linear and hierarchical dependencies in an artificial language. The second question relates to the factors influencing rule learning, namely the type of units, the role of phonological and prosodic cues in the inputs, and finally the contribution of semantic cue. In the next section we will outline the novelties characterizing our work compared to previous studies in artificial language learning.

Novelties of the present study

Compared to previous studies in the artificial language learning domain, we introduced three main novelties. First, the rules we investigated were closer to the ones we find in natural languages than the rules used in previous artificial language learning studies. Second, we addressed the question about the role of semantic cues in an artificial language learning study. Third, we investigated different types of knowledge and generalization. These three points are detailed in the following paragraphs.

1) As we mentioned above, the ALL studies exploiting the differences between Finite State Grammars and Phrase Structure Grammars led to different and often contradictory
conclusions. Thus, it seems that this research has reached a deadlock. For this reason, we proposed a new type of rule characterized by a hierarchical organization of the elements in the sentence. The rule used in Experiments 3 to 6 was designed to be closer to the rules we find in natural languages than the ones inspired by the Chomsky hierarchy. We exploited cues of different nature in order to implement this hierarchically based rule. We used phonological cues to indicate the grammatical category of the pseudo-words and the grammatical number, while we used prosodic cues to indicate the boundaries of the constituents. Therefore, in order to learn the subject-verb agreement rule, participants had to represent these two different levels of abstraction, on one hand the grammatical category and the number characterizing the pseudo-words, and on the other hand the structural relations.

2) Only a few studies (Moeser & Bregman, 1972; Morgan & Newport, 1981) directly tested the role of semantics in the learning of an artificial language. Moeser and Bregman (1972) argued that semantics was a necessary condition for the learning of complex grammatical relations. They showed that participants were able to learn their artificial grammar only when sequences of symbols were paired to the sentences cueing the syntactic structure. However, Morgan and Newport (1981) presented evidence that although semantics plays a role in the induction of complex syntactic systems it is not necessary. In Experiments 5 and 6, we directly investigate the role of semantic cues in the learning of the artificial language we created.

3) In artificial language learning studies, generalization has been investigated mainly as the transfer of the rule knowledge to a novel lexicon (e.g., Gomez & Gerken, 1999; Marcus et al., 1999). For instance, Gomez and Gerken (1999) showed that infants trained with word strings generated by a finite-state grammar (e.g., FIM SOG FIM FIM TUP) were able to generalize the grammar knowledge to other strings with the same pattern but with a novel lexicon (e.g., PEL TAM PEL PEL JIC). In our study, we not only tested this type of knowledge but we also tested the generalization of the rule knowledge to novel syntactic structures.

1.5.2. Empirical Framework

In order to study the learning of complex syntactic structures we created two rules of subject-verb agreement in number. We decided to use subject-verb agreement as it is
considered to be a prototypical example of structure-dependency. Indeed, in natural languages, the subject of a clause agrees with its verb even if they are not contiguous in the linear sequence. Subject and verb are linked in virtue of their position in the hierarchical configuration of the sentence. Therefore, agreement is a phenomenon that is computed on the basis of the hierarchical relations (i.e. c-command, locality) between the constituents of the sentence.

In the present dissertation we created two different subject-verb number agreement rules. The rule used in Experiments 1 and 2 was based on the position of the words in the linear sequence of the words, whereas the rule used in Experiments 3, 4, 5, and 6 relied on the hierarchical organization of the constituents in the sentence.

1.5.2.1. Types of Rule

a. Linear Subject-Verb Agreement Rule (Experiments 1 and 2)

The linear subject-verb agreement rule was defined as follows: the verb (V) agrees with the first noun (N) in the sentence. This agreement rule was based on the linear position of the elements (i.e., tone-words or pseudo-words) in a sentence. The words were combined to constitute three different types of sentence structures: $N_iV_i$, $N_iV_iN$, $N_iNV_i$ (in this notation the subscripts indicate the words involved in the agreement, the subject and the verb). As a result of the structures chosen, two types of syntactic dependencies were investigated: a) an adjacent dependency when the verb and the subject were contiguous in the structures, like in $N_iV_i$ and $N_iV_iN$; b) a non-adjacent dependency when there was an intervening noun between the subject and the verb, like in the structure $N_iNV_i$.

It is interesting to notice that the non-adjacent dependency studied in these experiments is quite similar to that usually employed in artificial language learning experiments. For instance, the dependency in the $N_iNV_i$ sentences (in which the first and the third units have the same number) is similar to the structures characterized by the AXB relationship in which the pseudo-words belonging to the category A were linked to the pseudo-words belonging to the category B (see for example Gomez, 2002; Onnis et al., 2003, 2004; Peña et al., 2002 among
others). This relationship as well as the linear subject-verb agreement rule is based on the linear structure of the sentence.

**b. Hierarchical Subject-Verb Agreement Rule (Experiments 3 to 6)**

The hierarchical subject-verb agreement rule was defined as follows: *the verb (V) agrees with the highest noun (N) in its constituent; if there is no noun in its constituent, the verb agrees with the highest noun in the immediately preceding constituent.* According to this rule, the agreement relation between the subject and the verb relied on the constituents’ structure of the sentence. In particular, in our experiments a prosodic cue (i.e., a pause) was used to mark constituent’s boundaries.

Three different types of sentence structures were created:

a) Two-word sentences consisted of a noun and a verb, N_iV_i; in this case the agreement rule led to an adjacent dependency.

b) Three-word sentences consisted of two nouns and a verb; in sentences N_iN_iV_i (where “_” indicates the position of the pause marking the constituent boundary) the agreement rule led to a non-adjacent dependency. In sentences N_i N_iV_i the agreement rule led to an adjacent dependency.

c) Four-word sentences consisted of two nouns and two verbs; in sentences N_iV_i N_iV_j the agreement rule led to two adjacent dependencies; in sentences N_iV_i N_iV_j the agreement rule led to both an adjacent and a non-adjacent dependency; both verbs agreed with the subject which was located at the beginning of the sentence.

A schematic representation of the different types of sentences is given in Figure 3. We wish to stress the fact that, since in our artificial language we only employed content words (nouns and verbs), our sentences were simplified compared to the similar sentence structures in natural language.
As already mentioned (see Section 1.4.2.1.2.), the studies that investigated the learnability of hierarchical structures used mainly hierarchical embedded structures (i.e., Phrase Structure Grammars). However, there is contradictory evidence about humans’ capacity to learn hierarchical embedded structures in artificial grammar paradigms (Fitch & Hauser, 2004; Friederici et al., 2006; Mueller et al., 2010; Perruchet & Rey, 2005; e.g., de Vries et al., 2011, 2008). We decided to create a new rule based on the constituents’ structure of the sentences. Phonological properties of the pseudo-words cued grammatical categories
(i.e., noun or verb) and number (i.e., singular or plural) while prosody cued the constituent structure.

1.5.2.2. Types of Knowledge

As we already pointed out in the review of the literature concerning ALL (see Section 1.4.2.4.), several types of generalization have been studied. The majority of ALL experiments explored infants and adults capacity to generalize their knowledge of the artificial language to a novel lexicon. On the contrary, little evidence has been collected about the generalization of the grammar knowledge to novel syntactic structures. In order to understand all aspects of the knowledge resulting from the learning of our artificial language, we decided to test different types of knowledge.

The characterization of these types of knowledge is the result of the comparison between the test materials and the training materials with particular attention to lexical items and syntactic structures. The four types of knowledge are represented schematically in Table 1.

Table 1

<table>
<thead>
<tr>
<th>Type of Knowledge</th>
<th>Lexical items</th>
<th>Syntactic structures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memorization</td>
<td>=</td>
<td>=</td>
</tr>
<tr>
<td>Lexical Generalization</td>
<td>≠</td>
<td>=</td>
</tr>
<tr>
<td>Syntactic Generalization</td>
<td>=</td>
<td>≠</td>
</tr>
<tr>
<td>Lexical and Syntactic Generalization</td>
<td>≠</td>
<td>≠</td>
</tr>
</tbody>
</table>
Memorization – This type of knowledge is investigated when participants perform the test phase with materials in which the *lexicon and the syntactic structures* are *identical* to the ones presented during the training phase.

Lexical Generalization – This type of knowledge is investigated when participants perform the test phase with materials in which the *syntactic structures* are *identical* to the ones presented during the training phase but the *lexicon* is constituted by *novel* pseudo-words.

Syntactic Generalization – This type of knowledge is investigated when participants perform the test phase with materials in which the *syntactic structures* are *novel* compared to the ones presented during the training phase but the *lexicon* is *identical*.

Lexical and Syntactic Generalization – This type of knowledge is investigated when participants perform the test phase with material in which the *lexicon and the syntactic structures* are *novel* compared to the ones presented during the training phase, but always consistent to the agreement rule.

1.5.3. Considerations about Data Analyses

The general assumption that guides the interpretation of artificial language learning paradigms’ results is that if participants perform significantly above chance they learned the grammar (or the rule) from the input. The most widely used measure for this kind of studies is *accuracy*, defined as the percentage of correct responses over the total number of items. This measure is then compared to chance level (e.g., the 50% for a task with two possible responses as it is the case for the grammaticality judgment task). However, a problematic aspect linked to the use of this measure is that accuracy does not take into account participant’s biases. We illustrate this point with an example. Imagine a participant performing a grammaticality judgment task. This participant has a bias toward the response “grammatical”. For instance, she answers “grammatical” to all grammatical items and to half of the ungrammatical items. This would result in accuracy above chance that does not reflect the learning of the grammar.

Alternative measures have been proposed in order to take into account participant’s bias. For instance, Perruchet and Pacteau (1990) proposed a measure called *$D$ scores*. This
measure is particularly suitable if the number of grammatical items is not the same as the number of ungrammatical items. This measure is defined as the percentage of ungrammatical items correctly categorized as “ungrammatical” minus the percentage of grammatical items incorrectly categorized as “ungrammatical”. D scores have been mainly used in studies in which the design of the paradigm required the presentation of a high number of ungrammatical items (Perruchet & Pacteau, 1990; Zimmerer, 2010; Zimmerer et al., 2010).

In the following section, we present another measure (i.e., d-prime) that is not often considered in artificial language learning studies, but may nevertheless represent a good way to characterize participants’ performance. D-prime has been extensively used in the Signal Detection Theory (Macmillan & Creelman, 1991) and only recently, it has also been adopted in artificial language learning studies (Flöel, de Vries, Scholz, Breitenstein, & Johansen-Berg, 2009; de Vries et al., 2010, 2008).

1.5.3.1. D-prime

D-prime is a measure that can be calculated for experiments with two possible answers (e.g., yes or no; grammatical or ungrammatical) and two classes of items (e.g., old or new; grammatical or ungrammatical). In the context of our experiments it represents an index of the ability to discriminate between grammatical and ungrammatical items. In other words, this measure represents an index of participants’ sensitivity to grammaticality. The use of this measure addresses the problem concerning participant’s biases. In fact, the justification for the use of $d$-prime as a measure of accuracy is that, as Macmillan and Creelman stated, “it is roughly invariant when response bias is manipulated; simpler indexes such as proportion correct do not have this property” (Macmillan & Creelman, 1991, p. 24).

In order to calculate these values participants’ responses are labelled as: Hit, False Alarm, Correct Rejection and Miss. A response is characterized as “Hit” when participants answer “grammatical” to a grammatical item, as “Correct Rejection” when participants answer “ungrammatical” to ungrammatical items, as “Miss” when participants answer “ungrammatical” to grammatical items, and as “False alarm” when participants answer “grammatical” to an ungrammatical item. Responses labels are schematized in Table 2.
Subsequently, in order to determine the d-prime values, the Hit Rate and the False Alarm Rate are calculated. These rates are defined as follows:

Hit Rate = proportion of grammatical items to which participant respond “grammatical”.

False Alarm Rate = proportion of ungrammatical items to which participant respond “grammatical”.

D-prime values are then calculated as:

\[ D-prime = Z(\text{Hit Rate}) - Z(\text{False Alarm Rate}) \] \[14\]

D-prime values can be positive or negative; the highest possible d-prime value is 6.93. A high d-prime value indicated high participant’s sensitivity and, consequently, a performance better than chance. A value of 0 means that grammatical and ungrammatical sentences are accepted at the same proportion, which means that the performance is at chance. As previously mentioned, this measure allows controlling for participants’ biases and response tendencies (e.g., the tendency to respond always “grammatical”). For example, a participant who applies the strategy of answering “grammatical” most of the times would probably reach a performance different than chance. She would then have a high Hit Rate but also a high False Alarm Rate. This would result in a low d-prime value which characterizes more efficiently than accuracy the sensitivity of the participant to the rule.

\[14\] In all experiments we calculated d-prime values using Pallier’s d-prime function implemented with R (Pallier, 2002):

\[
dprime <- \text{function} (\text{hit,fa}) \rightarrow \text{qnorm} (\text{hit}) - \text{qnorm} (\text{fa})
\]
In sum, in our experiments, *d-prime* values are calculated and used as dependent variable\(^{15}\). This measure allowed us to perform statistical analysis and a new type of analysis, namely the *Template Analysis*. As discussed below, the aim of the template analysis is to take into account the possibility that participants might have learned a rule different from the one they were supposed to learn from the input. Moreover, this analysis represents a valid instrument to investigate inter-participants’ performance variability.

**1.5.3.2. Template Analysis**

In artificial language learning paradigms it is often assumed that if participants answered correctly in a grammaticality judgment task they have learned the grammar, and otherwise they have not. In several studies concerning artificial language learning (e.g., Dulany et al., 1984; Endress et al., 2005; Pothos, Chater, & Ziori, 2006; Reber, 1967) a better-than-chance accuracy was interpreted as a confirmation that participants have learned a grammar (or a rule) from the input. The number of correct grammaticality judgments was then considered as a measure of the acquired knowledge of the grammar. However, other explanations of participants’ behaviour are possible (see Section 1.4.2.5.). In this context, an interesting distinction was introduced between “target grammar knowledge” and “fragmentary knowledge” (Perruchet & Pacteau, 1990; see also Pothos, 2007 for a review). The first term refers to the knowledge of the whole grammar actually used to generate the training items, and the second term refers to a partial memorization of the training items or their properties. Fragmentary knowledge may lead to a certain number of correct responses even if it is not derived from the target grammar. Therefore, Perruchet and Pacteau (1990) proposed that participants performed artificial grammar learning tasks comparing the similarity between new items and the items previously memorized during familiarization. Grammaticality judgements were then based on the degree of resemblance between the items used in the test and in the training. Dulany et al. (1984) developed another hypothesis to explain artificial grammar learning performances. They claimed that during training, participants were inferring a series of simple rules and not the target grammar. These “microrules” did not account for a perfect performance but could account for a better-than-chance performance. As a consequence of

\(^{15}\) All one-sample t-tests presented in this thesis will be one-tailed as we had a specific hypothesis about the direction of the difference between the sample mean and the value considered as chance (i.e., that the sample mean should be above chance and not simply different from chance).
these observations and the fact that there is usually high variability in group performances in ALL experiments, we believe that it is important to consider alternative rules participants might learn. Also Zimmerer (2010) pointed that a proper way to discuss artificial language learning experiments’ results involves the accurate investigation of individual performances and group heterogeneity.

It is possible that participants used alternative rules in order to accomplish the task which implies that their answers did not reflect the same rule representation. Exploring these alternative rules may lead to different interpretations of the results. The heterogeneity of the results, if properly considered, may provide new evidence for processes involved in artificial language learning. How could we examine which rule our participants were using during the grammaticality judgment task? In order to study which rule participants were using, we performed a template analysis. We defined a priori a set of potential alternative rules plausible considering the input given during the training, that we called “templates”. The aim of the templates analysis was to find a template which best characterized participants’ behaviour. To achieve this aim we analysed each set of data considering the fit of several possible templates for each experiment. The template with the highest d-prime value was considered as the most representative for the participant. A high d-prime, however, was not sufficient to characterize a performance as significantly different from chance. For this reason, we defined a performance as significantly above chance if the d-prime value was higher than 1. We decided to use this threshold as a d-prime of 1 corresponds to a percentage of correct responses of 69%, since a common threshold to indicate a performance better than chance is 66.66%, we considered a plausible solution to use of this threshold. The details of the description of the templates considered for each experiment are explained in the respective Results sections.

Caveat

A proper experimental approach requires that, when we compare two experiments, we only modify a single variable that may then be taken to explain the difference between the results. This approach allows experimenters to control what has been changed and to interpret the differences in the participants’ performance. Since the aim of this dissertation was mainly exploratory we decided to adopt a different approach. As a matter of fact, since the variables
that can be manipulated to create an artificial language are so numerous, the checking of the
effect of each single variable would have required a large number of experiments. Therefore,
we decided to manipulate more than one variable at a time trying to maximize the chances to
observe successful learning. We acknowledge that this choice will make the discussion of the
results more complex however we believe that will also allow us to cover a larger range of
phenomena.
Chapter 2. The Linear Syntactic Dependencies
2.1. Introduction

One aim of this thesis was to investigate rule learning in the linguistic domain and in a non-linguistic domain. The objective was to examine differences and similarities between the learning of regularities in language and in other domains, and to study cross-domain generalizations by means of an artificial language. We chose the music domain as the non-linguistic domain because of its similarity to language and because a number of studies have shown that learning regularities between tones is possible (Creel et al., 2004; Endress, 2010; Saffran et al., 1999). For instance, Saffran and colleagues (Saffran et al., 1999) investigated the segmentation process using a continuous stream of tones. Their participants, both adults and 8-month-old infants, were presented with a series of tones and tested on their capacity to extract word-like units relying only on statistical regularities between adjacent tones. The authors claimed that the same mechanism used to segment words in the language domain was used to extract units constituted by non-linguistic musical elements. Subsequently, Creel and colleagues (2004) investigated adults’ capacity to track statistical regularities among non-adjacent elements. They observed that this was possible when perceptual grouping cues (e.g., same pitch or timbre) were available.

Several studies have examined the parallels between language and music processing, especially looking for shared neural correlates (e.g., Jentschke, Koelsch, & Friederici, 2005; Koelsch, 2005; Koelsch, Gunter, Wittfoth, & Sammler, 2005; Patel, 1998, 2003, 2005). However, to our knowledge, no studies have directly explored the learning of grammars in the music domain by employing groups of tones as words to build up sentence-like sequences. Therefore, we chose to study the learning of a linear agreement rule using tone-words in Experiment 1. This rule is based on the linear position of the tone-words in the sequence. In this artificial language, tone-words were classified in categories that mimicked the grammatical categories “noun” and “verb” in natural language. In order to distinguish these two categories, a perceptual cue (i.e., different pitch) was employed. Each of these categories was further divided into two groups mimicking number (i.e., singular or plural). In order to distinguish singular and plural tone-words, a shorter extra-tone (the same for every word) was added as a suffix to the plural ones. Therefore, to successfully learn the agreement rule,
participants had to be able to categorize tone-words in the appropriate grammatical categories and number, and to discover that the verb agreed with the noun occupying the first position in the sentence. We thereby aimed to explore the possibility of learning syntactic-like patterns in a non-linguistic domain.

Subsequently, we wanted to compare participants’ performance in the musical domain with that in the linguistic domain. Therefore, we carried out experiment 2, in which the stimuli were pseudo-words. As in Experiment 1, this experiment introduced cues about the grammatical category (noun or verb) and number (singular or plural) of the pseudo-words. Grammatical category was determined by the manner of articulation of the consonants, while number was indicated by different final vowels. The comparison of the results of these two studies allowed us to evaluate the similarities and differences in regularity learning in a linguistic and a non-linguistic domain.
2.2. Experiment 1

In the present experiment the elements involved in the agreement rule were tone-words (a tone-word was built up of three subsequent piano tones), which were combined to form two and three tone-word sentences. It has been shown that infants and adults are able to extract tone-words from a continuous stream in a similar way as they do with linguistics stimuli (Saffran, 2002; Saffran et al., 1999). Creel and co-workers (2004) also showed that tracking non-adjacent relations among tones is more difficult than tracking adjacent relations. In fact, in their experiments, participants were not able to extract non-adjacent dependencies unless the elements involved in the dependency shared a perceptual cue (e.g., pitch or timbre), which differentiated them from the intervening element. This observation is in line with results showing that, also in the linguistic domain, non-adjacent dependencies are learned only under specific circumstances (e.g., Newport & Aslin, 2004; Peña et al., 2002).

In this artificial language we mimicked some properties of natural languages exploiting different types of cues. In particular, we categorized the tone-words in two groups that we called “grammatical categories”, since they were aimed to mimic the distinction between nouns and verbs that we find in natural languages. We implemented this distinction with pitch differences (high pitch for nouns and low pitch for verbs). In addition, in order to study number agreement we needed to differentiate singular from plural tone-words. Therefore we decided to mimic the plural marker by adding an extra tone, always the same, to the singular tone-word. This cue was inspired by languages like English or French in which number marking is implemented by the use of suffixes, such as the inflectional morpheme –s for English nouns. Therefore, in order to learn the linear agreement rule participants had to infer that the ‘verb’ (i.e., the low pitch tone-word) had to have the same number as the first ‘noun’ (i.e., the first high pitch tone-word) in the sentence (i.e., they both should either have or lack the extra tone). The main properties of the materials used in this experiment are summarized in Table 3.

16 Before choosing this type of number marker we studied in a pilot experiment the relevance of volume as perceptual cue. As participants did not pay attention at all to this cue to distinguish singular and plural tone-words we concluded that it was not salient enough.
In sum, with this experiment we aimed to add further observations to the debate about the learnability of regularities in a non-linguistic domain using a linear agreement rule. We expect participants to be able to rely on the cues marking grammatical categories and number to learn the rule and to learn both the adjacent and the non-adjacent dependencies showing a better-than-chance performance.

Table 3
Main properties of the materials used in Experiment 1.

<table>
<thead>
<tr>
<th>Stimuli Properties</th>
<th>Experiment 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary</td>
<td>Tri-tones</td>
</tr>
<tr>
<td>Grammatical Category Markers</td>
<td>Pitch</td>
</tr>
<tr>
<td>Nouns: high</td>
<td>Verbs: low</td>
</tr>
<tr>
<td>Number Markers</td>
<td>Suffix:</td>
</tr>
<tr>
<td>Singular: none</td>
<td>Plural: 165 ms G</td>
</tr>
<tr>
<td>Prosody</td>
<td></td>
</tr>
<tr>
<td>Pitch</td>
<td>No</td>
</tr>
<tr>
<td>Pauses</td>
<td>0ms between words</td>
</tr>
<tr>
<td>Agreement Rule</td>
<td>Linear</td>
</tr>
<tr>
<td>Locality</td>
<td>Adjacent and Non-adjacent</td>
</tr>
<tr>
<td>Match</td>
<td>Matching and Mismatching</td>
</tr>
<tr>
<td>Type of Knowledge</td>
<td>Memorization</td>
</tr>
</tbody>
</table>

2.2.1. Method

Participants

Eighteen participants took part in the experiment. They were all French speakers, with ages ranging between 19 and 40, with no reported hearing, reading or language impairment. They were paid or they received course credit for their participation.
Chapter 2. The Linear Syntactic Dependencies

Materials

Auditory Stimuli

Vocabulary: We constructed twelve piano tone-words. Each tone-word was composed of three single piano tones created with the software CuBase SX3 Steinberg. Each piano tone lasted 330 ms. These tone-words were similar to those used by Saffran and co-workers (Saffran et al., 1999). However, we introduced some changes as detailed below. Tone-words were characterized by different categories that mimicked the grammatical categories (i.e., noun and verb) and number (i.e., singular and plural) in natural language. The twelve words are listed in Table 4.

Table 4
Lexicon used in Experiment 1. Tone-words in bold font are characterized by high pitch. The smaller G at the end of the plural tone-words is the plural suffix.

<table>
<thead>
<tr>
<th>Grammatical Category</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ADB</td>
<td>ADBG</td>
</tr>
<tr>
<td></td>
<td>DFE</td>
<td>DFEg</td>
</tr>
<tr>
<td></td>
<td>GG#A</td>
<td>GG#Ag</td>
</tr>
<tr>
<td>Nouns</td>
<td>FCF#</td>
<td>FCF#g</td>
</tr>
<tr>
<td></td>
<td>D#ED</td>
<td>D#EDg</td>
</tr>
<tr>
<td></td>
<td>CC#D</td>
<td>CC#DG</td>
</tr>
<tr>
<td>Verbs</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Grammatical category marker: a different pitch implemented the distinction between grammatical categories, high pitch (i.e., one octave above the central one; from 523.15 Hz to 1000.6 Hz) for nouns and low pitch (i.e., one octave below the central one; from 108.56 Hz to 183.4 Hz) for verbs.

Number marker: the distinction between singular and plural tone-words was implemented by adding a suffix, namely a shorter tone (165ms G4) at the end of plural words. The same tone was used for plural nouns and plural verbs. This resulted in singular words constituted by three tones (total duration = 990 ms) and plural words constituted by four tones (total duration = 1155 ms).
Agreement rule: We created grammatical and ungrammatical sentences following the linear agreement rule (i.e., *the verb agrees with the first noun in the sentence*). As mentioned in section 1.5.2.1., three types of sentences were created: $N_iV_i$, $N_iV_iN_{ij}$, $N_iN_{ij}V_i$. The ungrammatical sentences were characterized only by the violation of the agreement rule (i.e., sentences with a singular subject and a plural verb or sentences with a plural subject and a singular verb). No position violations were introduced (i.e., there were no sentences in which the verb was in the first position).

Locality: If the subject and the verb were adjacent in the linear structure of the sentence, like in the structure NVN, the sentence was in the *Adjacent condition*; if there was an intervening element between subject and verb, like in the structure NNV, the sentence belonged to the *Non-adjacent condition*.

Match: the nouns in the three-word sentences either had the same number, in which case they were in the *Matching condition* (e.g., ADB DFE CC#D), or a different number, hence belonging to the *Mismatching condition* (e.g., ADB$_G$ DFE CC#D$_G$).

The complete set (test + training) of sentences included: 18 grammatical $N_iV_i$ sentences; 12 ungrammatical sentence $N_iV_j$; 72 grammatical $N_iN_jV_i$ sentences; 24 ungrammatical $N_iN_jV_j$ sentences; 72 grammatical $N_iV_iN_j$ sentences; 24 ungrammatical $N_iV_jN_j$ sentences. Examples of grammatical and ungrammatical sentences are given in Table 5. The sentences are reported in Appendix A.
Table 5

Examples of sentences used in Experiment 1. The subject and verb for each sentence are written in bold font. Ungrammatical sentences are marked with the symbol *.

<table>
<thead>
<tr>
<th>Sentence structure</th>
<th>Grammatical sentences</th>
<th>Ungrammatical sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>NV</td>
<td>DFE CC#D</td>
<td>*DFE CC#D</td>
</tr>
<tr>
<td></td>
<td>GG#A_G D#ED_G</td>
<td>*GG#A_G D#ED</td>
</tr>
<tr>
<td>NVN</td>
<td>ADB FCF# DFE_G</td>
<td>*ADB FCF#_G DFE_G</td>
</tr>
<tr>
<td></td>
<td>DFE_G D#ED_G ADB_G</td>
<td>*DFE_G D#ED ADB_G</td>
</tr>
<tr>
<td>NNV</td>
<td>ADB DFE_G FCF#</td>
<td>*ADB DFE_G FCF#_G</td>
</tr>
<tr>
<td></td>
<td>DFE_G ADC_G CC#D_G</td>
<td>*DFE_G ADC_G CC#D</td>
</tr>
</tbody>
</table>

Procedure

The experiment consisted of a Lexicon Familiarization, a Training phase, and a Test phase, as schematized in Figure 4. In the Lexicon Familiarization phase, participants were familiarized with the nouns and verbs of the lexicon through passive listening of all the tone-words (see below for details). The training and Test phases constituted the Rule Learning session. In the training phase, participants were asked to listen to a sample of sentences following the linear agreement rule and perform a detection task every time they listened to a plural tone-word. In the test phase, participants performed a grammaticality judgment task.

---

17 We decided to introduce the detection task in the training phase to maintain the participants’ attention. This was motivated by the fact that participants who had taken part in a previous pilot study had reported difficulty in attention maintenance during the training phase in which they had to listen passively to the sentences.
Lexicon Familiarization

Participants were presented with trials introducing nouns and verbs. Each trial had the following structure: a fixation cross was shown in the centre of the screen for 500 ms. Subsequently, the tone-word was presented auditorily followed by a 500 ms blank screen interval, which separated the trials. Participants were asked to listen to the tone-words carefully.

Rule Learning Session

Training phase: In this phase each trial had the following structure: a fixation cross was shown in the centre of the screen for 500 ms. A sentence was presented auditorily. Participants had to perform a detection task (pressing a central yellow button on the response box) every time they heard the shorter tone (G4) at the end of a tone-word. A 500 ms blank screen interval separated the trials. One hundred fourteen trials were presented.

Test phase: In this phase each trial had the following structure: a fixation cross was shown in the centre of the screen for 500 ms. Subsequently, the item was presented auditorily and the
response of the participant was recorded. If they judged that the sentence was grammatical, based on what they had heard in the training phase, they had to press the green button (otherwise the red button) on the response box. One hundred twenty trials were presented. Fifty percent of the sentences were grammatical.

### 2.2.2. Results

**D-Prime Analysis**

Before performing the statistical analyses, we eliminated the outliers. A response was defined as an outlier if the reaction time was 2.5 or more SD from the subject’s mean for each sentence’s length (2 tone-words; 3 tone-words). As a result of this procedure, 3.5% of the responses were dropped.

**Performance with respect to chance level**

Mean d-prime values for the matching and the mismatching conditions, for each type of knowledge and for each sentence structure tested in Experiment 1 are summarized in Table 6.

**Table 6.**

*Mean d-prime values (SD in parenthesis) for the Matching (M) and Mismatching (MM) conditions, for the type of knowledge and each sentence structure.*

<table>
<thead>
<tr>
<th>Type of Knowledge</th>
<th>NV</th>
<th>NVN</th>
<th>NNV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>MM</td>
<td>M</td>
</tr>
<tr>
<td>Memorization</td>
<td>0.31 (1.01)</td>
<td>0.62 (0.99)</td>
<td>0.20 (0.68)</td>
</tr>
</tbody>
</table>

In order to test the rule learning we performed several one-sample t-test comparing the mean d-prime values to chance level (i.e., zero for d-prime values).18

---

18 The use of t-tests is justified by the fact that the distribution of d-prime values meets the assumption of normality (p > .05 for Kolmogorov-Smirnov tests) for each type of structure.
Memorization – Performance was significantly above chance for the NVN structure in the matching condition ($t(17) = 2.650, p = .008$) and for the NNV structure in the matching condition ($t(17) = 3.039, p = .003$). Performance was not significantly above chance for the NV structure ($t(17) = 1.293, p = .106$) and for the NVN structure in the mismatching condition ($t(17) = 1.282, p = .108$); it was below chance for the NNV structure in the mismatching condition ($t(17) = -1.973, p = .032$).

Performance as a function of manipulated variables

We performed a 2x2 ANOVA on d-prime values in order to study the Locality (adjacent and non-adjacent location) and the Match (matching and mismatching condition) effects. In order to have the same number of observations for each condition, we considered only the three-word sentences. There was a significant main effect of the variable Match, $F(1,17) = 13.298, p = .002$. This effect reflected the fact that the performance of the participants was significantly better for the sentences in the matching condition ($M = 0.67, SD = 0.99$) that for the sentences in the mismatching condition ($M = -0.08, SD = 0.79$). The main effect of the variable Locality was marginally significant, $F(1,17) = 3.332, p = .086$; the performance of the participants was better for the sentences in the adjacent condition ($M = 0.41, SD = 0.86$) than for the sentences in the non-adjacent condition ($M = 0.18, SD = 1.06$). The interaction between Locality and Match was not significant, $F(1,17) = 2.123, p = .163$.

Template Analysis

We identified all templates that constituted plausible alternative rules the participants could have employed. The definition of the templates for this experiment is summarized in Table 7.
Table 7. 
*Description of the templates considered for Experiment 1.*

<table>
<thead>
<tr>
<th>Template</th>
<th>Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>1^-3^</td>
<td>The first and the third element have the same number</td>
</tr>
<tr>
<td>2^-3^</td>
<td>The second and the third element have the same number</td>
</tr>
<tr>
<td>1^-2^</td>
<td>The first and the second element have the same number</td>
</tr>
<tr>
<td>1^-2^-3^</td>
<td>The first, the second and the third elements have the same number</td>
</tr>
</tbody>
</table>

As shown in Figure 5, d-prime analysis showed that none of the participants learned the linear agreement rule. Four participants learned the template 1^-2^-3^. Fourteen participants did not reach the significance threshold for any of the other templates.

![Figure 5](image)

**Figure 5.** Individual participant performance as a function of the template with the highest d-prime value. Each circle represents a participant. The line represents the threshold above which a performance was considered to be significantly above chance.

**Reaction Times Analysis**

Table 8 summarizes the mean reaction times for correct responses for each type of sentence structure in matching and mismatching conditions and type of knowledge in Experiment 1.
Table 8.  
*Mean reaction times (SD in parenthesis) for each type of sentence structure, for Matching (M) and Mismatching (MM) conditions and type of knowledge.*

<table>
<thead>
<tr>
<th>Type of Knowledge</th>
<th>NV</th>
<th>NVN</th>
<th>NNV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>MM</td>
<td>M</td>
</tr>
<tr>
<td>Memorization</td>
<td>788</td>
<td>692</td>
<td>792</td>
</tr>
<tr>
<td></td>
<td>(298)</td>
<td>(213)</td>
<td>(293)</td>
</tr>
</tbody>
</table>

We performed an ANOVA on reaction times in order to study the Locality (adjacent vs non-adjacent) and the Match (match vs mismatch) effects. In order to have the same number of observations for each condition, we considered only the three-word sentences. There was a main effect of the variable Locality, $F(1,17) = 7.856$, $p = .012$. This effect reflected the fact that the adjacent dependency was judged faster than the non-adjacent dependency. The main effect of the Match variable was not significant, $F(1,17) = 1.826$, $p = .194$. There was a significant interaction between the variables, $F(1,17) = 5.883$, $p = .027$. The interaction reflected the fact that adjacent dependencies were judged faster than non-adjacent dependencies only for the matching condition ($t(17) = -3.780$, $p = .001$, two-tailed) and not for the mismatching condition ($t < 1$, *ns*).

2.2.3. Discussion

The aim of Experiment 1 was to investigate the learning of a linear agreement rule in the music domain. In this discussion we highlight three main results: 1) performance for the NV sentences was at chance; 2) there was a main effect of the Match variable; 3) four participants learned the template $1^\wedge-2^\wedge-3^\wedge$.

We observed that participants responded at chance in the grammaticality judgment task for the NV sentences. This result is particularly unexpected since these sentences were expected to be the easiest. To respond correctly to these sentences, participants only had to know that agreement is marked by the simultaneous presence (or absence) of a final identical tone on the two tone-words. This poor performance suggests that participants were unable to
use this number information to judge the grammaticality of the sentence and, consequently, that they did not learn the rule.

Second, we observed a main effect of the variable Match with better performance on the matching sentences than on the mismatching ones in the 3 tone-word sentences. In the matching grammatical sentences the three tone-words (those mimicking two nouns and a verb) all had the same number and thus the same final tone. As a result, the good performance observed in the matching sentences does not require participants to have learned the way grammatical categories were marked and/or to assign the syntactic role “subject” to the correct noun.

Third, in the template analysis, we observed that four participants learned the template 1^\text{-}2^\text{-}3^\text{\textsuperscript{\textcircled{\textasteriskcentered}}}, while the rest of the participants did not perform above chance for any template. We investigated whether we could relate the performance of these four participants to the Match effect, as it was possible that the better performance observed in the matching (vs. mismatching) sentences was entirely due to the performance of these four participants. We performed an ANOVA considering only the 14 participants who did not apply the 1^\text{-}2^\text{-}3^\text{\textsuperscript{\textcircled{\textasteriskcentered}}} template to examine if the Match effect was still significant or not. We observed that the main effect of the variable Match was still significant ($F(1,13) = 6.341, p = .026$). Thus, even the participants who did not learn this template judged matching sentences better than mismatching sentences. How can we explain this fact?

As previously mentioned, the matching sentences in the NNV and NVN grammatical conditions are perceptually salient, since all three tone-words in these conditions have the same grammatical number (i.e., they all either had or lacked the extra-tone). It is possible that even the participants who did not systematically apply the 1^\text{-}2^\text{-}3^\text{\textsuperscript{\textcircled{\textasteriskcentered}}} template benefited from this characteristic of the sentences. However, if this were the case, we should have observed more correct judgements not only for grammatical but also for ungrammatical matching sentences. Participants presumably should have rejected the ungrammatical matching sentences easily, since the verb had a different number compared to that of the two nouns. Since d-prime does not allow us to differentiate between grammatical and ungrammatical, we analyzed accuracy. As shown in Table 9, the above chance performance (70% accuracy) for
the matching sentences was restricted to the grammatical sentences; the ungrammatical ones were judged at chance (46% accuracy). Therefore, the explanation according to which participants based their judgments on the perceptual salience of the matching sentences cannot account for performance in the ungrammatical sentences, which should have been systematically rejected, but were not. In other words, we expected that participants who could use the 3 tone-word with matching final tones to judge grammatical matching sentences correctly could also infer that sentences without these 3 matching tones were ungrammatical. However, this was not the case. They appeared not to have learned the distinction between grammatical categories and performed at chance level both for the matching and the mismatching ungrammatical sentences. The fact that grammatical sentences were judged better than ungrammatical sentences could be due to memorization, since only the former but not the latter were presented in the training phase.

Table 9.
Mean accuracy values (SD in parenthesis) for the sentences in Experiment 1.

<table>
<thead>
<tr>
<th>Grammaticality</th>
<th>Matching</th>
<th>Mismatching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical</td>
<td>70 (20)</td>
<td>59 (21)</td>
</tr>
<tr>
<td>Ungrammatical</td>
<td>46 (22)</td>
<td>42 (19)</td>
</tr>
</tbody>
</table>

Given the poor performance of participants on the NV sentences and their systematically low d-prime scores on the other sentences types, we can only conclude that our participants did not learn the rule. However, it is more difficult to determine what is it they did learn and how they generated their responses in the grammaticality judgement task. We can hypothesize two possible scenarios. First, participants might have had a bias towards answering “grammatical”. This could explain the low d-prime values, as this measure takes into account participants’ biases, and also the better performance on grammatical matching sentences compared to the ungrammatical matching sentences. In order to test this hypothesis
we considered the number of “grammatical” and “ungrammatical” responses independently of whether the response was correct or not. We observed that the response “grammatical” represented the 58% of the total responses for three tone-word sentences, showing a tendency to be preferred over the response “ungrammatical”.

Another possible scenario could be that participants kept changing their response strategies during the test phase, trying to apply different rules without ever learning the correct one. This scenario is, however, difficult to investigate due to the high number of possible combinations of strategies participants might have adopted.

All in all, the present results suggest that participants were neither able to learn the linear agreement rule, nor to extract information about grammatical categories and number. Four explanations for this lack of learning are possible:

1) The use of the detection task during the training prevented learning. Although it forced participants to pay attention to the sentences throughout the whole session, the detection task might have directed their interest to single words in the sentences, thus preventing them from learning relations between the words. This explanation is purely speculative, as the data we possess do not permit us to test it.

2) The cues used to mark grammatical categories and number were not salient enough. However, this explanation would go against previous findings showing that adults are able to exploit pitch difference in ALL (Creel et al., 2004).

3) Keeping track of relations between tones and between tone-words was too difficult, as participants were not musically trained.

4) The linear agreement rule used in the present experiment cannot be learned in the music domain. In order to test this possible explanation, in the next experiment we investigated the learning of the same rule in the linguistic domain, trying to verify whether participants’ performance improves by changing the type of stimuli to be learned (musical vs linguistic).
2.3. Experiment 2

We performed Experiment 2 with linguistic stimuli, as we aimed to investigate the difference between regularities learning in a non-linguistic and in a linguistic domain. We tried to make the two experiments as similar as possible. We investigated the acquisition of the same agreement rule as in the previous experiment based on the linear position of the words in the sentence. The words involved in this rule were trisyllabic (e.g., patobu) which were combined to form two and three word sentences. We decided to use trisyllabic words because they matched the length of the three tone-words used in the previous experiment. Additionally, we wanted to extend the study of artificial language learning to the use of units longer than the monosyllabic words that have usually been utilized. To our knowledge, artificial language studies have only investigated the relationship between monosyllabic words (see for example Gomez, 2002; Marcus et al., 1999; Peña et al., 2002 among others).

In the present experiment the number markers were the final vowels. Grammatical categories were marked by phonological cues, as it has been shown that these cues represent an important source of information for identifying the grammatical categories of words (e.g., Shi et al., 1998; Shi, Werker, & Morgan, 1999)\(^{19}\). In order to learn the rule, participants had to learn that the verb agreed with the first noun in the sentence (i.e., both words had to share the same final vowel). The main properties of this experiment are summarized in Table 10.

\(^{19}\) This aspect of the materials, common to all the following experiments, will be detailed in the introduction of Chapter 3.
Table 10.
Main properties of the materials used in Experiment 2.

<table>
<thead>
<tr>
<th>Stimuli Properties</th>
<th>Experiment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary</td>
<td>Trisyllabic</td>
</tr>
<tr>
<td>Grammatical</td>
<td>Consonants manner of articulation:</td>
</tr>
<tr>
<td>Category</td>
<td>Nouns: plosive</td>
</tr>
<tr>
<td>Markers</td>
<td>Verbs: fricative</td>
</tr>
<tr>
<td>Number Markers</td>
<td>Final vowel:</td>
</tr>
<tr>
<td></td>
<td>Singular: /u/</td>
</tr>
<tr>
<td></td>
<td>Plural: /e/</td>
</tr>
<tr>
<td>Prosody</td>
<td>No</td>
</tr>
<tr>
<td>Pauses</td>
<td>0ms between words</td>
</tr>
<tr>
<td>Agreement</td>
<td>Linear</td>
</tr>
<tr>
<td>Rule</td>
<td>Adjacent and Non-adjacent</td>
</tr>
<tr>
<td>Match</td>
<td>Matching and Mismatching</td>
</tr>
<tr>
<td>Type of Knowledge</td>
<td>Memorization</td>
</tr>
</tbody>
</table>

We also modified the procedure with respect to the preceding experiment by introducing feedback during the grammaticality judgment task. We decided to use feedback in this experiment and the following ones because there is evidence that, in the context of an artificial language learning paradigm, feedback allows participants to reach a high level of competence in the artificial language (Dale & Christiansen, 2004). Even though the idea that it does not play a role in natural language learning is commonly accepted, studies about second language acquisition showed that adults exploit the presence of feedback to learn specific linguistic generalizations (see for example Carroll & Swain, 1993).

With the present experiment we aimed to investigate the role of the type of stimuli (the domain) in which a linear agreement rule is learned. We hypothesize that if there is an advantage of the linguistic domain over the music one for the learning of syntactic-like regularities, the rule should be learned better when linguistic stimuli are employed. If this hypothesis is correct, we expect participants’ performance in this experiment to be significantly above chance and also better than the performance observed in the music domain.
2.3.1. Method

Participants

Twenty-one participants took part in the experiment. They were all monolingual French speakers, with ages ranging between 19 and 25, with no reported hearing, reading or language impairment. They were paid or they received course credit for their participation.

Materials

Auditory Stimuli

Vocabulary: A lexicon of 12 trisyllabic (CV-CV-CV) words was created. The lexicon was constituted by three singular nouns and three plural nouns, three singular verbs and three plural verbs. The words were inspired by those used in Saffran et al., (Saffran, Newport, et al., 1996). The twelve words are listed in Table 11.

Grammatical category marker: the distinction between grammatical categories was implemented with the use of consonants with different manner of articulation, namely occlusive consonants [/p/, /t/, /b/] for nouns (e.g., patobu) and fricative consonants [/f/, /s/, /v/] for verbs (e.g., syvefu).

Number marker: the distinction between singular and plural words was implemented using different final vowels, namely /u/ for singular words and /e/ for plural words.

Table 11.
Lexicon used in Experiment 2 written in IPA.

<table>
<thead>
<tr>
<th>Grammatical Category</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nouns</td>
<td>patobu</td>
<td>patobê</td>
</tr>
<tr>
<td></td>
<td>têbipu</td>
<td>têbipê</td>
</tr>
<tr>
<td></td>
<td>bepêstu</td>
<td>bepêstê</td>
</tr>
<tr>
<td>Verbs</td>
<td>fôsevu</td>
<td>fôsevê</td>
</tr>
<tr>
<td></td>
<td>syvefu</td>
<td>syvefê</td>
</tr>
<tr>
<td></td>
<td>vofâsû</td>
<td>vofâsê</td>
</tr>
</tbody>
</table>
The words were recorded by a female native speaker of French using Cool Edit Pro 2.1 (22050Hz, 16 bit, Mono). The average intensity was set to 70 dB using Praat (Boersma & Weenink, 2007). The mean length of the words was 672 ms.

**Agreement rule:** The agreement rule investigated was the same as in Experiment 1 (i.e., *the verb (V) agrees with the first noun (N) in the sentence*).

**Locality:** As in Experiment 1, the rule led to sentences in the *Adjacent condition* and in the *Non-adjacent condition*.

**Match:** As in Experiment 1, sentences could be in the *Matching condition* (e.g., patobē syvefu tēbipu), or in the *Mismatching condition* (e.g., tēbipu patobē vofāsu).

The complete set (test + training) of sentences included: 18 grammatical N_iV_i sentences; 12 ungrammatical sentence N_iV_j; 72 grammatical N,N_jV_i sentences; 24 ungrammatical N,N_jV_j sentences; 72 grammatical N,V,N_j sentences; 24 ungrammatical N,V,N_j sentences. Examples of grammatical and ungrammatical sentences are given in Table 12. The sentences are reported in Appendix B.

**Table 12.**
*Examples of sentences used in Experiment 2 written in IPA. The subject and verb for each sentence are written in bold font. Ungrammatical sentences are marked with the symbol *.*

<table>
<thead>
<tr>
<th>Sentence structure</th>
<th>Grammatical sentences</th>
<th>Ungrammatical sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>NV</td>
<td>bepōtu fosevu</td>
<td>*bepōtu fosevē</td>
</tr>
<tr>
<td></td>
<td>patobē syvefē</td>
<td>*patobē syvefu</td>
</tr>
<tr>
<td>NVN</td>
<td>bepōtu fosevu</td>
<td>*bepōtu fosevē</td>
</tr>
<tr>
<td></td>
<td>tēbipu</td>
<td>tēbipu</td>
</tr>
<tr>
<td></td>
<td>vofāsē</td>
<td>vofāsu</td>
</tr>
<tr>
<td></td>
<td>patobu</td>
<td>patobu</td>
</tr>
<tr>
<td>NNV</td>
<td>bepōtu patobu</td>
<td>*bepōtu patobu</td>
</tr>
<tr>
<td></td>
<td>syvefu</td>
<td>syvefu</td>
</tr>
<tr>
<td></td>
<td>tēbipu</td>
<td>tēbipu</td>
</tr>
<tr>
<td></td>
<td>patobu vofāsē</td>
<td>*tēbipu patobu vofāsu</td>
</tr>
</tbody>
</table>

101
Procedure

The experiment consisted of a Lexicon Familiarization, a Training phase, and a Test phase like in Experiment 1, as schematized in Figure 6. The only differences were: in the training phase of the Rule Learning Session participants were asked to listen to a sample of sentences following the linear agreement rule without performing any detection task; participants received feedback during the test phase of the Rule Learning Session as explained below.

![Figure 6. Schematic representation of Experiment 2.](image)

Rule Learning Session

Test phase: In this phase each trial had the following structure: a fixation cross was shown in the centre of the screen for 500 ms. Subsequently, the item was presented and the response of the participant was recorded. If participants estimated that the sentence was grammatical on the grounds of what they had heard in the training phase, they had to press the green button (otherwise the red button) on the response box. A feedback (√ for correct answers and X for incorrect answers) lasting 1000 ms was then presented. A 500 ms blank screen interval
separated the trials. One hundred twenty trials were presented. Fifty percent of the sentences were grammatical. Accuracy and reaction times were recorded.

2.3.2. Results

D-Prime Analysis

Before performing the statistical analyses, we eliminated the outliers. A response was defined as an outlier if the reaction time was 2.5 or more SD from the subject’s mean for each sentence’s length (2 words; 3 words). As a result of this procedure, 2.7% of the responses were dropped.

Performance with respect to chance level

Mean d-prime values for each type of sentence structure for matching and mismatching conditions and type of knowledge investigated are summarized in Table 13.

Table 13.
Mean d-prime values (SD in parenthesis) for the Matching (M) and Mismatching (MM) conditions, for each type of knowledge and sentence structure.

<table>
<thead>
<tr>
<th>Type of Knowledge</th>
<th>NV</th>
<th>NVN</th>
<th>NNV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memorization</td>
<td>2.17 (1.89)</td>
<td>0.97 (1.96)</td>
<td>0.26 (1.10)</td>
</tr>
</tbody>
</table>

In order to test the rule learning we performed several one-sample t-test comparing the mean d-prime values to the chance level (i.e., zero is the chance level for d-prime values).

Memorization – Performance was significantly above chance for the NV structure ($t(20) = 5.241, p < .001$). Performance was significantly above chance for the NVN structure in the matching condition ($t(20) = 2.275, p = .017$) and for the NNV structure also in the matching condition ($t(20) = 3.017, p = .003$). Performance was marginally above chance for

20The use of t-tests is justified by the fact that the distribution of d-prime values meets the assumption of normality (p>.05 for Kolmogorov-Smirnov tests) for each type of structure.
the NVN structure in the mismatching condition ($t(20) = 1.082, p = .073$) and not significantly above chance for the NNV structure in the mismatching condition ($t < 1, ns$).

**Performance as a function of manipulated variables**

We performed an ANOVA on d-prime values in order to study the Locality and the Match effects. In order to have the same number of observation for each condition we considered only the three word sentences. There was a significant main effect of the variable Match, $F(1,20) = 6.033, p = .023)$. This effect reflected the fact that the performance of the participants was significantly better for the sentences in the matching condition ($M = 1.11, SD = 1.91$) that for the sentences in the mismatching condition ($M = 0.21, SD = 1.22$). The main effect of the variable Locality and the interaction between Locality and Match were not significant ($Fs < 1$).

**Template Analysis**

The templates considered for this experiment were the same as in Experiment 1. As shown in Figure 7, d-prime analysis showed that none of the participants learned the linear agreement rule. Nine participants learned the template $1^\wedge-2^\wedge-3^\wedge$. The other 12 participants did not reach the significance threshold for any of the templates.
Figure 7. Individual participant performance as a function of the template with the highest d-prime value. Each circle represents a participant. The line represents the threshold above which a performance was considered as significantly above chance.

Reaction Times Analysis

Mean reaction times for correct responses for each type of sentences structure for matching and mismatching conditions and type of knowledge investigated are summarized in Table 14.

Table 14. Mean RTs (SD in parenthesis) for the Matching (M) and Mismatching (MM) conditions, for the type of knowledge and each sentence structure.

<table>
<thead>
<tr>
<th>Type of Knowledge</th>
<th>NV</th>
<th>NVN</th>
<th>NNV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>MM</td>
<td>M</td>
</tr>
<tr>
<td>Memorization</td>
<td>1201</td>
<td>870</td>
<td>1025</td>
</tr>
<tr>
<td></td>
<td>(339)</td>
<td>(343)</td>
<td>(385)</td>
</tr>
</tbody>
</table>
We performed an ANOVA on reaction times in order to study the Locality and the Match effects. In order to have the same number of observation for each condition, we considered only the three word sentences. There was a main effect of the Match variable, $F(1,20) = 14.278$, $p = .001$. This effect reflected the fact that sentences in the matching condition were judged faster ($M = 872$ ms, $SD = 336$) than sentences in the mismatching condition ($M = 1003$ ms, $SD = 352$). The main effect of the variable Locality and the interaction were not significant ($Fs < 1$).

2.3.3. Discussion

The aim of Experiment 2 was to study the acquisition of the linear agreement rule in the linguistic domain. The results of the present experiment and those of Experiment 1 allow us to compare in different domains the learning of an agreement rule based on the linear position of the elements in the sentence.

In contrast with Experiment 1, the participants’ performance to NV sentences here was significantly above chance. This observation suggests that participants at least learned how agreement was marked, that is, that the agreeing words must share the same final vowel. Since there were only two words in the sentence, it was not necessary to distinguish the grammatical categories or to take into account the position of the subject and the verb in the linear sequence of the words. Therefore, we can conclude that participants generally managed to understand the way number was implemented and agreement was marked.

Two other important results emerged: 1) a main effect of the variable Match; 2) a large number of the participants (i.e., 9 out of 21 participants) used the template 1^\-2\^\-3\ above chance. Once again, these two results may be connected in the sense that the Match effect may be attributable to the performance of participants using this template. To verify this we performed a mixed ANOVA on d-prime values. The between-subject variable was the Group, that is whether a participant belongs to the group using the 1^\-2\^\-3\ template or to the group constituted by the rest of the participants. The within-subject variables were Locality and Match. As expected, the analysis revealed a main effect of the variable Match ($F(1,19) = 27.172$, $p < .001$): participants performed better in the matching sentences
Chapter 2. The Linear Syntactic Dependencies

\( M = 1.33, SD = 1.91 \) than in the mismatching ones \( (M = 0.24, SD = 1.22) \). Also the main effect of the Group variable was significant \( (F(1,19) = 27.172, p < .001) \). Participants belonging to the group using the template performed better \( (M = 1.67, SD = 1.96) \) than the rest of participants \( (M = -0.09, SD = 0.79) \). The main effect of Locality was not significant \( (F < 1) \). Most interestingly, the interaction between the Match and Group variables was significant \( (F(1,19) = 42.906, p < .001) \). This interaction reflected the fact that the difference between the matching and mismatching was significant only for the group using the template \( (t(8) = 7.341, p < .001) \) and not for the rest of participants \( (t(11) = -1.071, p = .307) \). None of the other interactions was significant \( (Fs < 1) \). We can conclude that the Match effect observed was due to those participants who applied the template. Therefore, it appears as if we can identify two groups: a first one with the nine participants who learned the template 1\(^{\text{-}}\)2\(^{\text{-}}\)3\(^{\text{-}}\); and a second with the twelve participants who did not learn the agreement rule or any of the alternative templates. Participants belonging to the latter group performed at chance level on all the types of sentence structure\(^{21} \). Therefore, we assume that they did not apply any rule or learn any aspect of the artificial language.

Concerning the participants who learned the alternative template, we can ask why they learned a different rule. It is possible that during training they focused only on grammatical matching sentences, since they are perceptually more salient than grammatical mismatching sentences. As a matter of fact, both grammatical matching NVN and NNV sentences have the same sequence of final vowels (u-u-u for singular; ė-ē-ē for plural). This pattern can presumably be perceived easily on the sole basis of these rhymes. The presence of such a pattern might have prevented participants from learning the correct linear agreement rule and facilitated the learning of the incorrect alternative template. If participants indeed applied this rule, we can conclude that they did not differentiate grammatical categories or pay attention to words’ linear position in the sentence.

To sum up, the results of Experiment 2 suggest that none of the participants learned the linear agreement rule. Participants could be divided into two groups according to their

\(^{21} \) We performed additional one-sample t-test to compare the performance of this group of participants to chance level. The results indicated that they performed at chance for all types of three-word sentence structures (NVN matching: \( t(11) = -1.405, p = .094 \); NVN mismatching, NNV matching and NNV mismatching: \( t < 1, ns \)).
performance. One group performed at chance level in all conditions and did not learn the linear agreement rule or any of the alternative templates. The second group of participants learn the template 1^ - 2^ - 3^ based on the perceptual salience of the grammatical matching sentences, but did not learn the linear agreement rule.

2.4. Comparison between Experiment 1 and Experiment 2

2.4.1. Results

D-prime Analysis

Mean d-prime values for each sentence structure, matching condition and experiment are represented in Figure 8.

![Figure 8](image)

**Figure 8.** Mean d-prime values for Experiment 1 and Experiment 2 (Error bars represent half standard deviation).

We performed a mixed ANOVA on the d-prime values. The between subject variable was the *Experiment* (i.e., Experiment 1, Experiment 2); the two within subject variables were
Locality and Match. The analysis revealed a main effect of Match: $F(1,37) = 14.218, p = .001$. Performance was better for matching sentences ($M = 1.11, SD = 1.91$) than for mismatching sentences ($M = 0.21, SD = 1.22$). There was no significant main effect of Locality ($F < 1$) or Experiment ($F(1,37) = 1.520, p = .225$). The interaction between Locality and Match was marginally significant ($F(1,37) = 2.911, p = .096$). This interaction reflected the fact that the Match effect was significant both for the adjacent and the non-adjacent dependencies (Adjacent dependency: $t(38) = 2.013, p = .051$; Non-adjacent dependency: $t(39) = 4.539, p < .001$), but was bigger for the non-adjacent dependency (matching-mismatching difference = 1.09) than for the adjacent dependency (matching-mismatching difference = 0.58). There were no other significant interactions (Match and Experiment, Locality and Match and Experiment, $Fs < 1$).

Reaction Times Analysis

We performed a mixed ANOVA on the reaction times. The between subject variable was the Experiment (i.e., Experiment 1, Experiment 2); the two within subject variables were Locality and Match. The analysis revealed a main effect of Match: $F(1,37) = 13.264, p = .001$. Participants were faster in Experiment 1 ($M = 770$ ms, $SD = 259$) than in Experiment 2 ($M = 938$ ms, $SD = 349$). The main effect of Locality was not significant ($F < 1$). The interaction between Locality and Experiment was marginally significant: $F(1,37) = 3.434, p = .072$. The difference between adjacent and non-adjacent dependencies was significant for Experiment 1 ($t(17) = -2.803, p = .012$), but not for Experiment 2 ($t < 1, ns$). Also the interaction between Match and Experiment was marginally significant: $F(1,37) = 3.304, p = .077$. The difference between matching and mismatching conditions was significant for Experiment 2 ($t(20) = -3.779, p = .001$), but not for Experiment 1 ($t(17) = 1.351, p = .194$). The analysis revealed that also the interaction between Locality and Match was significant: $F(1,37) = 4.212, p = .047$. The difference between matching and mismatching conditions was significant for the adjacent dependency ($t(38) = -4.406, p < .001$) but not for the non-adjacent dependency ($t(38) = -1.559, p = .127$). Finally the interaction between Locality, Match and Experiment was not significant ($F < 1$).
2.4.2. Discussion

The aim of the comparison of these two experiments was to investigate the differences in the learning of a linear agreement rule in the music domain and in the linguistic domain. The results of the two experiments revealed that participants did not learn the rule in either domain.

In both experiments there was an effect of the variable Match. However, the template analyses suggested different origins for this effect. We assumed that in Experiment 1, participants had a bias towards the answer “grammatical” which explains the low d-prime values and the performance at chance, or they kept changing their answer strategy. In experiment 2 we distinguished two groups of participants, one applying the template 1^2^3^ and one answering at chance levels for all three-word sentences.

Since participants did not learn the rule in either experiment, we cannot draw any final conclusions about the comparison of the musical and the linguistic domains. It is possible that investigating a linear agreement rule was an injudicious choice. In fact, as we do not find such type of rule in natural language, learning could have been difficult. We took advantage of the negative results of the present experiments to change the agreement rule investigated. In the next experiments we decided to focus only on the linguistic domain and to examine the learning of a hierarchical agreement rule closer to the syntactic rules we can find in natural languages. Several changes will be introduced in the next experiments.

The most important change is that the new rule we will investigate is conditioned by the hierarchical structure of the sentences of the artificial language\textsuperscript{22}. Hierarchy will be marked by prosodic cues as detailed in the next section. The relevance of several phonological cues to the learning will be also explored. We concluded that the results of the first two experiments, and especially of Experiment 2, were due to a perceptual bias triggered by the fact that all words (nouns and verbs) with the same number had the same final vowel. This identity relation on which the agreement relationship was based, and the presence of the

\textsuperscript{22} The hierarchical subject-verb agreement rule was defined as follows: the verb (V) agrees with the highest noun (N) in its constituent; if there is no noun in its constituent, the verb agrees with the highest noun in the immediately preceding constituent (see Section 1.5.2.1.).
matching condition, could have favoured the perceptual bias development. Therefore, we decided that in the next experiment the agreement relationship will be based on the co-occurrence of perceptually not identical markers. In order to better control the characteristics of the stimuli, we will synthesize the sentences using a natural prosody and manipulating the presence of pauses, instead of simply recording the words and presenting them in sequences as we did in the previous experiment. Moreover, we decided to extend the investigation of rule learning testing also to different types of knowledge. In particular, we tested participants’ capacity to learn the rule and generalize it to a new lexicon and/or to novel syntactic structures.
Chapter 3. The Hierarchical Syntactic Dependencies
3.1. Introduction

In the previous chapter we discussed the learning of a linear agreement rule in two different domains, namely music and language. We uncovered the difficulties of learning syntactic dependencies using non-linguistic material and then decided to focus only on the language domain. In the present chapter we present three experiments which aim is to investigate the learning of a hierarchical agreement rule. We used this rule because, as outlined in section 1.5.1., it is more similar to those we find in natural language than the rule used in Experiments 1 and 2. The hierarchical agreement rule was defined as a relationship between elements belonging to different grammatical categories, namely nouns and verbs, which occupy specific positions in the hierarchical structure of the sentence, as subject and verb. In order to learn the rule, three learning steps by the participant are necessary:

1) Learning the grammatical categories and number (i.e., learning that there are two grammatical categories of words. Learning that words can be either singular or plural).
2) Learning the morphological agreement markers and their association (i.e., learning the co-occurrence of the markers).
3) Learning the hierarchical agreement rule: learning to parse the sentence into constituents, to identify the syntactic roles (i.e., subject and verb), and learning the agreement between the subject and the verb (e.g., learning that in the sentence N_NV the subject is the noun in the same constituent as the verb, while in the sentence NN_V the subject is the noun in the preceding constituent).

In Experiments 3, 4 and 5, the information about grammatical categories and number was carried by various phonological cues (i.e., vowel and consonant features, number of syllables and syllable structure), while the information about constituent structure was provided by prosodic cues (i.e., pitch contour and pauses).

We decided to use phonological cues to mark grammatical category and number based on the observation that the linguistic input that language learners receive contains phonological cues and these learners rely on them to acquire the lexicon and the syntax of their language.
(Morgan & Demuth, 1996b). As Kelly (1996) pointed out, there are several phonological cues such as syllable duration, vowel quality and consonant quality that correlate with grammatical categories. By the first year of life, infants acquire the capacity to distinguish and exploit these characteristics. For instance, in a series of studies Shi and co-workers showed that newborns are already able to use perceptual cues, like the ones mentioned above, to discriminate between words of different form classes (Shi et al., 1998, 1999). In particular, these studies revealed that acoustic and phonological cues are available in the input and are sufficient to allow newborns to distinguish, for instance, between content and function words. Therefore, phonological cues represent an important source of information to categorize words, which is exploited already very early in language acquisition. Concerning number marking in natural languages, we can observe that in some languages, like English and French, number is characterized by recurring suffixes. English and French plural nouns are regularly marked with the ending morpheme –s (e.g., in French more than 98% of plural adjectives and nouns end in –s). In the following experiments, we chose to convey the same information as suffixes, namely to indicate if a word is singular or plural by means of phonological cues such as consonant and vowel features of the final phonemes.

We decided to use prosodic cues to mark constituent boundaries as it has been shown that syntactic and prosodic boundaries often coincide (Fisher & Tokura, 1996; Truckenbrodt, 1999). As Cooper & Paccia-Cooper (1980) observed, pauses tend to occur more between phrases than within a phrase. This results in the fact that prosody plays a role in syntactic parsing in adults (Cutler, Dahan, & van Donselaar, 1997; Eckstein & Friederici, 2005, 2006). In addition, psycholinguistic studies showed that prosodic cues such as pauses and pitch changes are used in first language acquisition. Infants are not only sensitive to this kind of cues (Gerken et al., 1994; Jusczyk, Hohne, & Mandel, 1995), but they are also able to use them to recognize clauses (Nazzi, Kemler Nelson, Jusczyk, & Jusczyk, 2000) and phrases (Soderstrom, Seidl, Kemler Nelson, & Jusczyk, 2003) in continuous speech. For instance, Soderstrom and colleagues (Soderstrom et al., 2003) showed that 6- and 9-month-old infants displayed preference for passages containing well-formed verb phrases over passages containing units crossing subject and verb phrase, when strong prosodic cues marked the boundaries of the two types of units.
Furthermore, in a few artificial language studies, the influence of prosodic cues on word segmentation has been investigated (see for example Bagou et al., 2002; Peña et al., 2002; Shukla et al., 2007). These studies showed that the presence of prosodic markers allows participants to extract word-like sequences from continuous speech. Interestingly, Morgan and colleagues (Morgan et al., 1987) claimed that, although prosodic cues do not indicate the entire hierarchical structure of sentences, they mark the main constituents, and therefore they play a role in language acquisition. Employing an artificial language, they demonstrated that the presence of prosodic cues marking phrase boundaries facilitated the learning of a finite state grammar. Thus, we decided to use prosodic cues to mark constituent boundaries as it has been shown that they play a role in artificial language processing facilitating the acquisition of a grammar. This will allow us to extend the investigation of the role of prosodic cues to the learning of a hierarchical agreement rule. In order to introduce these cues we decided to apply a natural pitch contour to synthesized sentences with pseudo-words and to manipulate the length of the pauses marking constituent boundaries.

As already mentioned in section 1.5.1., the second aim of this dissertation was to investigate the relevance of different factors influencing rule learning. For this reason, we systematically manipulated phonological and prosodic cues in the following experiments. In particular, in Experiments 3 and 4, we used vowel and consonant features, and we also manipulated the variability of syllables not carrying agreement cues, as it will be detailed in the introduction of the experiments. In Experiment 5 we manipulated the number of syllables and syllable structure to mark grammatical category and the final vowels to indicate the number. In addition, we increased the salience of the pause at constituents’ boundaries across the experiments and added pauses between the words. We thereby investigated the importance of prosody in the characterization of the syntactic structures.

To sum up, the general aim of the following experiments was to investigate the learning of a hierarchical agreement rule while the second aim was to investigate the relevance of different kinds of phonological and prosodic cues employed on rule learning. The characteristics and differences between experiments will be detailed in the introduction of each experiment.
3.2. Experiment 3

In Experiment 3, we investigated the learning of a hierarchical agreement rule. In this rule, the agreement relation between the subject and the verb relies on the constituent structure of the sentence and not on the constituents’ linear position in the word string. Bisyllabic pseudo-words constituted the lexicon. Grammatical categories and number information were implemented by phonological cues (i.e., features of consonants and vowels), while prosodic cues (i.e., pauses) marked constituency. The place of articulation of the consonants (i.e., bilabial vs. alveolar) in the second syllable was used to mark the grammatical category (i.e., verb vs. noun) and their voicing feature (i.e., voiced vs. unvoiced) cued number (i.e., singular vs. plural). Grammatical categories were marked also by the height of the vowels in the second syllable (i.e., close vowels for nouns and open vowels for verbs), and number was marked by vowels’ backness (i.e., back vowels for singular and front vowels for plural). Crossing all the phonological cues we obtained four possible second syllables (i.e., -du, -bo, -ti, -pa).

The first syllable (i.e., the stem) of the words belonged to a large set (i.e., 92 different syllables), in line with the high variability of word stems in natural languages. The role of the first syllables was to create variability in the signal in contrast to the second syllables, which constituted the less variable part of the signal, critically involved in the rule. The high variation of the stem aimed to drive participants’ attention to the second one. As Gomez (2002) observed, when adults and infants perform a task in which they have to learn syntactic dependencies of the type AXC, they tend to focus primarily on adjacent dependencies. On the contrary, if there is high variability of the X element (i.e., 16 - 24 different Xs), participants are able to extract the non-adjacent dependency between the A and C elements. In contrast to the material used by Gomez, in the present experiment the position of the highly variable element is within the words and not between the words implicated in the relationship. However, the role of variability is the same, namely to create noise in order to highlight the recurrent elements involved in the syntactic dependency. We assumed that since the first syllables belong to a large set, participants would focus on the relationships between second syllables (i.e., the suffixes). Therefore, they would be able to exploit the phonological cues characterizing grammatical category and number.
Once participants detect these elements (i.e., morphological agreement markers) and the association between them, they would possess the building blocks necessary to learn the rule. However, in order to learn the hierarchical agreement rule participants should be able to parse the sentences into constituents exploiting the presence of the pauses marking constituents’ boundaries. They have to learn the syntactic roles and to identify the position of the subject, which varies according to the hierarchical structure of the sentence rather than to the linear order of the words. Moreover, the effects of three variables characterizing the materials were studied, namely Locality, Match and Type of knowledge. With the term Locality we refer to the fact that the hierarchical agreement rule leads to two different types of dependencies, namely adjacent dependencies when the subject and the verb belong to the same constituent (like in sentence N_N_i V_i), and non-adjacent dependencies when they belong to different constituents, like in the sentence N_i N_V_i. The Match variable refers to the fact that the two nouns in the sentences either had the same number (i.e., matching condition) or different number (i.e., mismatching condition). Finally, the Type of knowledge variable refers to types of knowledge investigated in the experiment as a result of the differences and similarities between the characteristics of the test and training materials (see Section 1.5.2.2.).

In the present experiment two types of knowledge were investigated: Lexical Generalization and Lexical and Syntactic Generalization. The former was studied when the vocabulary used in the test phase was different than the one used in the training phase; the latter was explored when both the vocabulary and the syntactic structures presented in the test phases were different from the ones used in the training phase.

To sum up, in the present experiment participants were trained with sentences made of bisyllabic words. They had to learn a hierarchical agreement rule exploiting the cues available in the signal, namely phonological cues in the suffixes of each word and pauses between constituents. The main properties of this experiment are summarized in Table 15. Given the resemblance of this artificial grammar to natural language grammar, we expected participants to be able to learn the rule by showing significantly above chance performance.
Table 15.  
**Main properties of the materials used in Experiment 3**

<table>
<thead>
<tr>
<th>Stimuli Properties</th>
<th>Experiment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary</td>
<td>Bisyllabic: CCV-CV</td>
</tr>
<tr>
<td>Grammatical Category Markers</td>
<td>2nd syllable: Nouns: Alveolar consonant Close vowel Verbs: Bilabial consonant Open vowel</td>
</tr>
<tr>
<td>Number Markers</td>
<td>2nd syllable: Singular: Voiced consonant Back vowel Plural: Unvoiced consonant Front vowel</td>
</tr>
<tr>
<td>Prosody</td>
<td>Pitch: Natural Pauses: 0ms between words 100ms between constituents</td>
</tr>
<tr>
<td>Agreement Rule</td>
<td>Hierarchical</td>
</tr>
<tr>
<td>Locality</td>
<td>Adjacent and Non-adjacent</td>
</tr>
<tr>
<td>Match</td>
<td>Matching and Mismatching</td>
</tr>
<tr>
<td>Type of Knowledge</td>
<td>Lexical Generalization Lexical and Syntactic Generalization</td>
</tr>
</tbody>
</table>

3.2.1. Method

**Participants**

Twenty participants took part in the experiment. They were all monolingual French speakers, with ages ranging between 19 and 45 with no reported hearing, reading or language impairment. They were paid or they received course credit for their participation.

**Materials**

*Vocabulary:* A lexicon of 192 bisyllabic words was created. All words had a first syllable with a structure CCV (e.g., bla-, ste-, etc.) and a second syllable with a structure CV (i.e., -du, -bo, -ti or -pa). In order to create these words, 96 CCV syllables were chosen, all respecting the phonotactic constraints of French. Half of these syllables (i.e., 48 syllables) were used to create nouns and half to create verbs. Each first CCV syllable was used for the singular and the plural forms of the word. Thus each first syllable was represented twice in the lexicon.
Examples of words are listed in Table 16. Words were characterized by grammatical categories (i.e., noun and verb) and number (i.e., singular and plural); this information was marked by phonological cues in the second syllable as detailed below.

**Grammatical category markers:** The place of articulation of the consonants in the second syllable (CV) marked the grammatical category; final syllables for nouns contained alveolar consonants (i.e., /t/ and /d/), while final syllables for verbs contained bilabial consonants (i.e., /p/ and /b/). The difference between grammatical categories was also implemented by the height of the vowels in the second syllables; closed vowels (i.e., /u/ and /i/) characterized nouns, while mid-open and open vowels (i.e., /o/ and /a/) characterized verbs.

**Number markers:** Differences in voicing were used to distinguish the number; second syllables for singular words were characterized by voiced consonants (i.e., /b/ and /d/) while second syllables for plural words were characterized by voiceless consonants (i.e., /p/ and /t/). The difference between singular and plural was also implemented by vowel backness in the second syllables; back vowels (i.e., /u/ and /o/) characterized singular words and front vowels (i.e., /i/ and /a/) characterized plural words.

**Table 16.**

Examples of words used in Experiment 3. Consonants characterizing the grammatical category and number are in italic font, while the final vowels characterizing the grammatical category and number are in bold font.

<table>
<thead>
<tr>
<th>Grammatical Category</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nouns</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>bla-du</td>
<td>bla-ri</td>
</tr>
<tr>
<td></td>
<td>ste-du</td>
<td>ste-ri</td>
</tr>
<tr>
<td></td>
<td>kRo-du</td>
<td>kRo-ri</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td><strong>Verbs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>bRo-bo</td>
<td>bRo-pa</td>
</tr>
<tr>
<td></td>
<td>fla-bo</td>
<td>fla-pa</td>
</tr>
<tr>
<td></td>
<td>sle-bo</td>
<td>sle-pa</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

23 Words listed in the text and in tables are always written in SAMPA alphabet.
Prosody: Words were then combined to build up sentences. In order to apply a natural prosody to these sentences a trained female French speaker recorded French sentences with the same number of phonemes and the same syntactic structure as the artificial sentences, respecting the natural French prosody. The pitch contour of the sentences was then extracted using Praat (Boersma & Weenink, 2007) and applied with the French diphone (fr1) of MBROLA speech synthesizer (Dutoit, Pagel, Bataille, & Vreken, 1996) to the sentences of the experiment. A 100 ms pause was added in order to mark the constituents’ boundaries.

Agreement Rule: As mentioned in section 1.5.2.1, we created five different types of syntactic structures combining the words in the sentences. The grammatical sentences followed the hierarchical agreement rule (i.e., the verb (V) agrees with the highest noun (N) in its constituent; if there is no noun in its constituent, the verb agrees with the highest noun in the immediately preceding constituent). Ungrammatical sentences were characterized by number agreement violations (i.e., sentences with a singular subject and a plural verb or sentences with a plural subject and a singular verb). No position violations were introduced (i.e., there were no sentences in which the verb was in the first position). Examples of grammatical and ungrammatical sentences are given in Table 17.

**Table 17.**

*Examples of sentences used in Experiment 3. The subject and verb of each sentence are in bold. Ungrammatical sentences are marked with the symbol *.*

<table>
<thead>
<tr>
<th>Sentence structure</th>
<th>Grammatical sentences</th>
<th>Ungrammatical sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>NV</td>
<td>blidu slibo vRiti kRipa</td>
<td>*vlOdu stapa fREti bRibo</td>
</tr>
<tr>
<td>N_NV</td>
<td>fRati _ kRudu gREbo pRati _ stEtI blepa</td>
<td>*kladu _ pludu vlupa fRadu _ kleti pRybo</td>
</tr>
<tr>
<td>NN_V</td>
<td>glydu sIOti _ fRubo klati stoti _ slipa</td>
<td>*dREdu sIOdu _ vRypa blyti dReti _ kREbo</td>
</tr>
<tr>
<td>NV_NV</td>
<td>glEdu kRebo _ pRidu pRubo stedu kRybo _ dROti glapa</td>
<td>*kRodu glapa _ pROdu pRupa klydu bRupa _ fRyti pRubo</td>
</tr>
<tr>
<td>NVN_V</td>
<td>flOdu slebo steti _ glabo pRiti bRupa bRoti _ flEpa</td>
<td>*bRodu vRapa pRidu _ glapa flOti plybo klydu _ flEbo</td>
</tr>
</tbody>
</table>
Locality: The hierarchical agreement rule led to different types of dependencies, namely adjacent dependency and non-adjacent dependency. If the subject and the verb were adjacent in the hierarchical structure of the sentence, like in structure N_N_i V_i, the sentence was in the Adjacent condition; if there was an intervening word between the subject and the verb, like in N_i N_V_i, the sentence was in the Non-adjacent condition.

Match: The nouns in the three and four-word sentences either had the same number, in which case the sentence was in the Matching condition (e.g., klati stoti _ slipa), or a different number, in which case the sentence was in the Mismatching condition (e.g., fRati _ kRudu gRebo).

Type of knowledge: In the present experiment we investigated Lexical Generalization and Lexical and Syntactic Generalization. Lexical Generalization was investigated when the lexicon used in the test phase was different than the one used in the training phase; the Lexical and Syntactic Generalization was explored when both the vocabulary and the syntactic structures used in the test phases were different from the ones used in the training phase.

The sentences used in this experiment included: 84 grammatical two-word sentences (NV); 28 ungrammatical two-word sentence; 280 grammatical three-word sentences (140 sentences N_NV and 140 sentences NN_V); 104 ungrammatical three-word sentences (52 sentences N_NV and 52 sentences NN_V); 56 grammatical four-word sentences (28 sentences NV_NV and 28 NVN_V); 56 ungrammatical four-word sentences (28 sentences NV_NV and 28 NVN_V). The sentences are listed in Appendix C.

Procedure

The experiment consisted of training and test phases as schematized in Figure 9. In the Lexicon Familiarization phase, the words were presented auditorily and in isolation for the participants to become familiar with them; in the Rule Learning phase they were asked to passively listen to sentences following the hierarchical agreement rule and to perform a grammaticality judgment task. The sentences were presented in random order.
Lexicon Familiarization

Participants were presented with trials introducing the words. Each trial had the following structure: a fixation cross was shown at the centre of the screen for 500 ms; subsequently, the word was presented auditorily and followed by a 300 ms blank screen interval which separated the trials. All words were presented once. Participants were asked to listen to the words carefully.

Rule Learning Session

This session was constituted by two series of training and test phases followed by a last test phase without training. In the first series, NV sentences with two-words were presented (48 sentences in the training and 48 sentences in the test); in the second series, NV and NNV sentences with two and three-words were presented (184 sentences in the training phase and
184 sentences in the test phase); the last test phase was constituted by NNV and NVNV sentences with three and four-words (144 sentences).

**Training phases:** In these phases each trial had the following structure: a fixation cross was shown at the centre of the screen for 500 ms. A sentence was then acoustically presented. A 500 ms blank screen interval separated the trials. Participants were asked to listen to the sentences carefully.

**Test phases:** In these phases each trial had the following structure: a fixation cross was shown at the centre of the screen for 500 ms. A sentence was presented auditorily and participants had to judge whether it was grammatical or not according to what they had heard in the training phase. Their choice was communicated by pressing the green or red button on the response box. A feedback (√ for correct answers and X for incorrect answers) lasting 1000 ms was then presented and a 500 ms blank screen interval separated the trials. Fifty percent of the sentences were grammatical. Once participants reached 80% accuracy, they automatically moved to the next training phase; if they did not reach that threshold, they kept performing the task until the end of the phase.

### 3.2.2. Results

**D-prime Analysis**

Before performing the statistical analyses we eliminated outliers. A response was classified as an outlier if the reaction time was two or more SD from the subject’s mean for each sentence length (two-words; three-word; four-word). As a result of this procedure, 5% of the responses were dropped.

**Performance with respect to chance level**

Mean d-prime values for the matching and the mismatching conditions, for each Type of knowledge and for each sentence structure tested in this experiment are summarized in Table 17.
Table 17.  
Mean d-prime values (SD in parenthesis) the Matching (M) and Mismatching (MM) conditions, for each Type of knowledge and sentence structure.

<table>
<thead>
<tr>
<th>Type of Knowledge</th>
<th>NV</th>
<th>N_NV</th>
<th>NN_V</th>
<th>NV_NV</th>
<th>NVN_V</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>MM</td>
<td>M</td>
<td>MM</td>
<td>M</td>
</tr>
<tr>
<td>Lexical Generalization</td>
<td>1.06</td>
<td>1.78</td>
<td>-0.03</td>
<td>1.92</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>(1.56)</td>
<td>(2.31)</td>
<td>(0.99)</td>
<td>(2.48)</td>
<td>(1.09)</td>
</tr>
<tr>
<td>Lexical and Syntactic Generalization</td>
<td>1.50</td>
<td>0.64</td>
<td>1.65</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.61)</td>
<td>(1.61)</td>
<td>(2.44)</td>
<td>(1.17)</td>
<td></td>
</tr>
</tbody>
</table>

In order to test if participants performance was above chance we performed several one-sample t-tests comparing the mean d-prime values to the chance level (i.e., as previously mentioned, zero is the chance level for d-prime values)\textsuperscript{24}.

**Lexical Generalization** – Performance was significantly above chance for the NV structure ($t(19) = 3.056, p = .003$), for the N_NV structure in the matching condition ($t(19) = 3.441, p = .001$) and for the NN_V structure in the matching condition ($t(19) = 3.463, p = .001$). Performance was marginally significant for the NN_V structure in the mismatching condition ($t(19) = 1.417, p = .086$) and not significantly different from chance for the N_NV structure in the mismatching condition ($t < 1, ns$).

**Lexical and Syntactic Generalization** – Performance was significantly above chance for the NV_NV structure in the matching condition ($t(19) = 2.568, p = .009$) and for the NVN_V structure in the matching condition ($t(19) = 3.034, p = .003$). Performance was also significantly above chance for the NV_NV structure in the mismatching condition ($t(19) = 1.797; p = .044$), but not significant for the NVN_V structure in the mismatching condition ($t < 1, ns$).

\textsuperscript{24} The use of t-tests is justified by the fact that the distribution of d-prime values meets the assumption of normality ($p > .05$ Kolmogorov-Smirnov tests) for each type of structure.
Performance as a function of manipulated variables

We performed an ANOVA on d-prime values in order to study the effect of the Type of knowledge, Locality and Match. Only the Match variable had a significant main effect, \( F(1,19) = 10.808, p = 0.004 \), with the effects of Type of knowledge and Locality being non significant (\( F < 1 \)). The former effect reflected the fact that performance was significantly better for the matching sentences (\( M = 1.71, SD = 2.41 \)) than for the mismatching sentences (\( M = 0.29, SD = 1.23 \)). There was also a significant interaction between Type of knowledge and Match, \( F(1,19) = 4.133, p = 0.056 \). The effect of the Match was significant for both types of knowledge (Lexical Generalization: \( t(19) = 3.409, p = .003 \); Lexical and Syntactic Generalization: \( t(19) = 2.873, p = .010 \)), but it was bigger for the Lexical Generalization (Matching-mismatching difference = 1.69) than for the Lexical and Syntactic Generalization (Matching-mismatching difference = 1.16) as can be observed in Figure 10. There were no other significant interactions (Type of knowledge and Locality: \( F(1,19) = 1.907, p = .183 \); Locality and Match: \( F < 1 \); and Type of knowledge, Locality and Match: \( F(1,19) = 1.449, p = .243 \)).

![Figure 10](image-url)

**Figure 10.** Mean d-prime values for the Matching and Mismatching conditions for Lexical Generalization and Lexical and Syntactic Generalization Types of knowledge.
Template Analysis

As already explained in section 1.5.3.2., all templates were plausible (defined a priori) alternative rules which might have been used by participants. We considered different templates for sentences with three and four words. The definition of the templates for this experiment is summarized in Table 18.

Table 18.
Description of the templates considered for Experiments 3.

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Template</th>
<th>Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>Three Pseudo-words</td>
<td>1^-3^</td>
<td>The first and the third pseudo-words have the same number</td>
</tr>
<tr>
<td></td>
<td>2^-3^</td>
<td>The second and the third pseudo-words have the same number</td>
</tr>
<tr>
<td></td>
<td>NpauseV</td>
<td>The noun before the pause and the verb have the same number</td>
</tr>
<tr>
<td>Four Pseudo-words</td>
<td>1^-4^</td>
<td>The first and the fourth pseudo-words have the same number</td>
</tr>
<tr>
<td></td>
<td>3^-4^</td>
<td>The third and the fourth pseudo-words have the same number</td>
</tr>
</tbody>
</table>

Three-Word Sentences

As shown in Figure 11, two participants had the highest d-prime value for the hierarchical subject-verb agreement rule and they performed above chance. Three participants performed above chance with the first template and two participants with the second template. The performance of thirteen participants was not above chance for any of the above mentioned templates.
Figure 11. Individual participant performance for the three-word sentences as a function of the template with the highest d-prime value. The line represents the threshold above which performance was considered as above chance.

Four-Word Sentences

As shown in Figure 12, four participants had the highest d-prime value for the hierarchical subject-verb agreement rule and they performed above chance. Four participants performed above chance with the third template. The performance of twelve participants was not above chance for any of the templates.
Chapter 3. The Hierarchical Syntactic Dependencies

Figure 12. Individual participant performance for four-word sentences as a function of the template with the highest $d$-prime value. The line represents the threshold above which performance was considered as above chance.

**Reaction Times Analysis**

Mean RTs for correct responses for the Matching and the Mismatching conditions, for each Type of knowledge and for each sentence structure tested are summarized in Table 19.

**Table 19.**

*Mean RTs (SD in parenthesis) in ms for correct responses for the Matching (M) and Mismatching (MM) conditions, for Type of knowledge and sentence structure.*

<table>
<thead>
<tr>
<th>Type of Knowledge</th>
<th>NV</th>
<th>N_NV</th>
<th>NN_V</th>
<th>NV_NV</th>
<th>NVN_V</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>MM</td>
<td>M</td>
<td>MM</td>
<td>M</td>
</tr>
<tr>
<td>Lexical Generalization</td>
<td>732</td>
<td>723</td>
<td>815</td>
<td>683</td>
<td>820</td>
</tr>
<tr>
<td></td>
<td>(241)</td>
<td>(280)</td>
<td>(363)</td>
<td>(260)</td>
<td>(348)</td>
</tr>
<tr>
<td>Lexical and Syntactic Generalization</td>
<td>549</td>
<td>587</td>
<td>635</td>
<td>673</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(236)</td>
<td>(274)</td>
<td>(257)</td>
<td>(338)</td>
<td></td>
</tr>
</tbody>
</table>
We performed an ANOVA on reaction times in order to study the Type of knowledge, the Locality and the Match effects. There was a significant main effect of the variable Type of knowledge, $F(1,19) = 29.424$, $p < .001$. This effect reflected the fact that participants were faster in the Lexical and Syntactic Generalization ($M = 611$ ms, $SD = 278$) than in the Lexical Generalization ($M = 760$ ms, $SD = 315$). There was a significant main effect of the variable Locality, $F(1,19) = 4.348$, $p = .051$; participants were faster in judging adjacent dependencies ($M = 669$ ms, $SD = 306$) than non-adjacent dependencies ($M = 703$ ms, $SD = 306$). There was also a significant main effect of the Match variable, $F(1,19) = 6.039$, $p = .024$; participants were faster in the matching condition ($M = 648$ ms, $SD = 262$) than in the mismatching condition ($M = 724$ ms, $SD = 341$). There were two significant interactions. The interaction between Type of knowledge and Locality ($F(1,19) = 6.883$, $p = .017$) was significant, with adjacent dependencies judged faster than non-adjacent dependencies only for the Lexical and Syntactic Generalization ($t(19) = -3.479$, $p = .003$), while there was no significant difference between adjacent and non-adjacent dependencies for the Lexical Generalization ($t < 1, ns$). The significant interaction between Type of knowledge and Match variables ($F(1,19) = 4.747$, $p = .042$) reflected the fact that matching sentences were judged faster than mismatching sentences for the Lexical Generalization ($t(19) = -2.640$, $p = .016$), but not for the Lexical and Syntactic Generalization ($t(19) = -1.483$, $p = .155$). The interaction between Locality and Match ($F(1,19) = 1.1242$, $p = .302$) and between Type of knowledge, Locality and Match ($F(1,19) = 1.266$, $p = .275$) were not significant.

3.2.3. Discussion

The main purpose of the present experiment was to investigate whether adult participants were able to learn a hierarchical agreement rule in an artificial language learning paradigm. In order to better understand the results of the present experiment, we shall first recall that there are three learning steps that participants had to go through to learn the hierarchical agreement rule. The first step is the identification of grammatical categories and number. The second step concerns the learning of the morphological agreement markers and their association. In the present experiment the morphological agreement was determined by the association between the second syllables (i.e., suffixes) of the words; in order to learn this association, participants had to focus on second syllable co-occurrence (i.e., a word ending
with the syllable –bo must go with a word ending with the syllable –du, and a word ending with the syllable –pa must go with a word ending with the syllable –ti). In the last step participants learn the hierarchical agreement rule which is required to parse the sentence into constituents and to identify the syntactic roles.

We begin our discussion by considering the first two steps. We can test if participants attained these steps by looking at their performance to the matching sentences. Our assumption is that in order to judge these sentences participants only needed to identify the grammatical categories and number and to identify the morphological agreement. It was not necessary to be able to parse the sentences into constituents and to identify the syntactic roles, as the verb agreed with both nouns in the grammatical sentences and it disagreed with both nouns in the ungrammatical sentences. Therefore, to perform the grammaticality judgment task correctly, participants did not need to reach step 3. Considering the results of the t-tests comparing participants’ performance for matching sentences to the chance level, we observe that this performance is significantly above chance for every type of sentence structure considered (i.e., N_NV, NN_V, NV_NV, NVN_V). Moreover, there is a main effect of Match which indicates that matching sentences are always judged better than mismatching sentences. Two strategies concerning the learning of grammatical categories and number are possible to explain these results. On the one hand, participants might have identified the two grammatical categories (i.e., noun and verb) and number (i.e., singular and plural) as assumed in step 1. On the other hand, it is possible that they identified four distinct categories and based the associations on those. From our data it is not possible to disentangle these two alternatives. However, the results did show that participants are able to focus on the relevant part of the words (i.e., second syllables) and that they managed to identify the correct associations between syllables. This observation suggests that - whatever the strategy employed by the participants - it enabled them to identify the morphological agreement markers and their associations correctly, and reach step 2.

Did our participants reach step 3 and learn the hierarchical agreement rule? Since the analysis of the performance to the matching sentences does not allow us to disentangle step 2 from step 3, we consider only the mismatching condition. We observe from the results of the t-tests comparing participants’ performance for mismatching sentences to the chance level.
(i.e., four comparisons), that only for one comparison the performance is significantly above chance (i.e., NV_NV), for one comparison it is only marginally above chance (i.e., NN_V) and for two other comparisons (i.e., N_NV and NVN_V) performance is not significantly different from chance. These results suggest that participants did not reach step 3 and did not learn the hierarchical agreement rule. Interestingly, there is also an interaction between Match and Type of knowledge. This interaction reflects the fact that the difference between matching and mismatching is reduced for the Lexical and Syntactic Generalization as compared to the Lexical Generalization. This is possible because performance to the mismatching condition for the Lexical and Syntactic Generalization improves thanks to the NV_NV structure. The NV_NV structure could be processed as two independent NV structures. As shown by the performance on the NV sentences, participants were able to correctly judge this type of structure, and therefore they could have applied the same principle also to the NV_NV structure. However, the knowledge of grammatical categories, number, and the association between morphological agreement markers (i.e., step 1 and 2) is sufficient to judge this type of sentence without parsing the constituent structure. In addition, the peculiar status of the NV_NV structure –the fact that it could be easier than other structures– is confirmed by reaction times. Participants were faster in judging adjacent dependencies than non-adjacent dependencies only in the Lexical and Syntactic Generalization. For the Lexical and Syntactic Generalization, the adjacent dependencies correspond to the NV_NV sentences.

The template analysis suggests that the matching condition might have prevented participants from learning the correct rule. This analysis reveals that some participants seem to have learned a linear rule. As Endress and co-workers pointed out (Endress et al., 2009) repetitions and similarity play an important role in artificial grammar learning. The presence of these perceptual primitives might have influenced the type of regularities participants can learn. In the present experiment, the sentences in the matching condition could have contributed to the erroneous learning of a linear rule by deviating participants’ attention to properties of the input (i.e., the identity between the nouns, du-du or ti-ti) that are irrelevant to the hierarchical agreement rule. We also observe that in the matching condition, the answers for the hierarchical agreement rule and for the templates converge (i.e., when a sentence was grammatical for the hierarchical agreement rule, it was grammatical for the templates; when a
sentence was ungrammatical for the hierarchical agreement rule it was ungrammatical for the templates). This resulted in the fact that, even if participants were applying an incorrect rule (i.e., one of the alternative templates) during the grammaticality judgment task, in the matching condition they were always receiving positive feedback. This could have prevented participants from learning the correct hierarchical agreement rule, as they were positively reinforced in the use of an incorrect rule.

If the matching condition indeed was less informative than the mismatching condition and it probably had a negative influence on learning, why did we introduce it? In the original goals of this experiment we wanted to draw a parallel between subject-verb agreement in natural and artificial languages. The general idea was to compare error rates in the artificial language learning paradigm with attraction errors in natural language paradigms\(^{25}\). We wanted to investigate in an artificial language whether an intervening element would attract agreement differently, depending on its nature and its position in the hierarchical structure of the sentence. Even though the tasks used in these paradigms are of course quite different (i.e., production task, sentence reading, and grammaticality judgment task), we could have obtained interesting correspondences. In order to perform this comparison, we needed an intervening element of the same nature as the subject (e.g., singular when the subject is singular) or of a different nature (e.g., plural when the subject is singular). This led to the creation of the matching and mismatching conditions. In our original paradigm, the presence of the intervening element would have triggered more errors in the mismatching condition than in the matching condition based on its position in the hierarchical structure of the sentence. However, because participants did not learn the rule, the comparisons between matching and mismatching conditions were not informative.

In light of the results presented here and in order to improve rule learning in the following experiment, we decided to make three changes. First, we eliminated the matching condition to avoid the negative effects this might have had on learning. In Experiment 4, the two nouns in the same sentence are always different in number (i.e., one is singular and one is plural). Second, we decided to completely eliminate variability in the signal (i.e., the stems) in

\(^{25}\) In natural language production paradigms, attraction errors are defined as subject-verb agreement errors in which the verb agrees with a word, often called “intervening element”, that is not the subject of the sentence.
the next experiment, which resulted in a lexicon made up of six words. This choice was made in order to minimize lexical processing and promote syntactic processing. The third change concerns the addition of a pause (i.e., 25 ms) between the words. It is possible that in the present experiment, participants were forced to perform two cognitive tasks at the same time; on one hand they were segmenting the signal in words and on the other hand they were learning the hierarchical agreement rule. To overcome this problem and favour the rule learning task over the word segmentation task, we decided to add a pause between words. This choice was inspired by a study from Peña and colleagues (2002). They showed that adults were able to extract grammatical higher level structures from speech only when a short pause marking the units’ boundaries was available in the signal. If the pause was not introduced in the signal, they were only able to segment the pseudo-words without extracting the rule characterizing them. These modifications will be detailed in the introduction of the next experiment.

3.3. Experiment 4

The results of the previous experiment showed that participants were not able to successfully learn the hierarchical agreement rule implemented in our artificial language, even though they seemed to identify grammatical categories, number and the morphological markers association. In light of the discussion about the factors influencing this result, we decided to make some changes in the artificial language used in the present experiment. The aim of Experiment 4 is to study whether adults improved their ability to learn a hierarchical agreement rule under these new conditions.

Three main changes were made compared to Experiment 3. The first modification concerns the inclusion of the matching condition. In the previous experiment we observed that including the matching condition created difficulties in the results and could have prevented participants from learning the correct agreement rule. As a matter of fact, participants only needed to identify grammatical categories and number, to learn the morphological agreement markers and their association, in order to be able to correctly judge the sentences belonging to the matching condition. We suggested that the matching condition might have had a negative effect on learning, since participants could receive positive feedback applying rules that were
not the correct one. Therefore, in the present experiment we decided to include only the *mismatching condition*.

The second change relates to the variability of the word stems, that is, elements that were not involved in the syntactic dependency (i.e., the first syllable of the words). In Experiment 3, we exploited the variability of the first syllable of the words with the aim of directing the attention of the participants to the invariant part of the words (i.e., the second syllable), which was critically involved in the rule. This idea was implemented using a large number of first syllables and only four possible second syllables to build up the words. However, variability may have increased the lexical processing load, and therefore, make the learning of the rule more difficult. In the present experiment we decided to simplify the lexical processing in order to favour syntactic processing. To do this, we eliminated the first syllables and reduced the lexicon to four monosyllabic words (i.e., du, bo, ti, pa) constituted by the second syllables of the words used in the previous experiment. Therefore the vocabulary was constituted by a much *reduced set of monosyllabic words*. The vocabulary was completed with two additional monosyllabic words introduced to test Lexical Generalization (see below for further details).

The third modification relates to words segmentation. We assumed that one of the reasons why participants did not learn the hierarchical agreement rule in Experiment 3 was that they were performing two tasks at the same time: segmenting the acoustic signal into words and extracting regularities. In the present experiment, we chose to add a 25 ms pause between the words. This choice intended to facilitate segmentation and to ensure that more cognitive resources were available and dedicated to rule extraction. As Peña and colleagues showed in their experiments (Peña et al., 2002), participants were able to extract higher level grammatical structures only with the presence of a short *segmentation cue* (i.e., a 25 ms pause); if the pause was not available, they were able to segment the pseudo-words, but not to extract the rule characterizing them. As a result, we added a 25 ms pause between the words and we lengthened the pauses between constituents to 125 ms in order to have pauses of different length for segmentation and to mark the constituents’ boundaries. The main properties of the present experiment are summarized in Table 20.
Table 20.  
Main properties of the materials used in Experiment 4.

<table>
<thead>
<tr>
<th>Stimuli Properties</th>
<th>Experiment 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary</td>
<td>Monosyllabic: CV</td>
</tr>
<tr>
<td>Grammatical</td>
<td>Consonants:</td>
</tr>
<tr>
<td>Category</td>
<td>Nouns: Alveolar and velar</td>
</tr>
<tr>
<td>Markers</td>
<td>Verbs: Bilabial</td>
</tr>
<tr>
<td>Number Markers</td>
<td>Consonants:</td>
</tr>
<tr>
<td></td>
<td>Singular: voiced</td>
</tr>
<tr>
<td></td>
<td>Plural: unvoiced</td>
</tr>
<tr>
<td>Prosody</td>
<td>Natural</td>
</tr>
<tr>
<td>Pitch</td>
<td>25ms between words</td>
</tr>
<tr>
<td></td>
<td>125ms between constituents</td>
</tr>
<tr>
<td>Pauses</td>
<td>Hierarchical</td>
</tr>
<tr>
<td>Agreement</td>
<td>Adjacent and Non-adjacent</td>
</tr>
<tr>
<td>Rule</td>
<td>Mismatching</td>
</tr>
<tr>
<td>Locality</td>
<td>Memorization</td>
</tr>
<tr>
<td>Match</td>
<td>Lexical Generalization</td>
</tr>
<tr>
<td>Type of Knowledge</td>
<td>Syntactic Generalization</td>
</tr>
<tr>
<td></td>
<td>Lexical and Syntactic Generalization</td>
</tr>
</tbody>
</table>

In the present experiment, we decided to complete the study of rule learning and rule generalization introducing the investigation of two other types of knowledge: Memorization and Syntactic Generalization. As a result, in Experiment 4 we investigated all four types of knowledge mentioned in section 1.5.2.2., namely Memorization, Syntactic Generalization, Lexical Generalization, and Lexical and Syntactic Generalization.

Memorization refers to an item-based type of knowledge, since we examine it when both the lexicon and syntactic structures used in the test phases are identical to the ones used in the training phases. Syntactic Generalization is involved when the lexicon is the same in the training and in the test phases but participants are tested on new syntactic structures (they were trained on three-word sentences and they are tested on four-word sentences). In contrast to Experiment 3, in which it was possible to create new words by introducing new stems (i.e., first syllables) while keeping the same suffixes (i.e., second syllables), testing generalization to a new lexicon in Experiment 4 required the creation of nouns with new suffixes. Two novel nouns (ge and kE) were introduced in a lexicon familiarization session as well as in a training
phase involving NV sentences (with the previously learned verbs). Hence, participants should learn that the new noun ge belongs to the same category as the noun du and the noun kE belongs to the same category as the noun ti. Participants were then tested on three-word sentences. As a result, Lexical Generalization in Experiment 4 differs from that in Experiment 3 because it relies on the introduction of novel words that do not share the suffix with the previously learned lexicon, as was the case in Experiment 3. Participants had to infer whether a word belonged to a certain grammatical category and number only on the basis of the training with the novel NV sentences without any perceptual similarity between the members of the same category. In a sense, this type of knowledge is more similar to a morphological generalization than to a lexical generalization, since the former involves new words and suffixes, whereas the latter involves identical suffixes appended to variable word stems. However, since in order to demonstrate their knowledge that the same rule applies to new words, in both cases participants have to learn that the words belonging to the same grammatical category and number are interchangeable, we decided to use the terminology Lexical Generalization for both experiments.

To sum up, in the present experiment participants are presented with the same artificial language as in Experiment 3, with a novel lexicon constituted by monosyllabic words. We expected that with the suppression of the matching condition, the simplified lexicon and the presence of pauses between words, participants would be able to learn the hierarchical agreement rule, that is their performance would be significantly above chance.

3.3.1. Method

Participants

Thirty participants took part in the experiment. They were all monolingual French speakers, with ages ranging between 20 and 45, with no reported hearing, reading or language impairment. They were paid or they received course credit for their participation.

---

26 It is interesting to note that although some languages like Italian have a limited set of nominal number suffixes, others like French have a much more variable set of suffixes (at least in spoken language). Hence, the two different ways by which we implemented nominal number features in Experiments 3 and 4 are actually found in natural languages.
Materials

Vocabulary: A lexicon of six monosyllabic words was created. Four of these words were constituted by the second syllables (i.e., du, bo, ti, pa) used in Experiment 3 and two new nouns (i.e., ge, kE) were created. Words are listed in Table 21. Words were characterized by grammatical category (i.e., noun and verb) and number (i.e., singular and plural); this information was marked by phonological cues as detailed below.

Grammatical category marker: The place of articulation of the consonants indicated the grammatical category; two alveolar consonants (i.e., /t/ and /d/) were employed for nouns used to study Memorization and Syntactic Generalization, two velar consonants (i.e., /g/ and /k/) were employed for nouns used to study Lexical Generalization and Lexical and Syntactic Generalization. Verbs were characterized by two bilabial consonants (/p/ and /b/).

Number marker: Voicing distinguished the number of the word; singular words were characterized by voiced consonants, while plural words were characterized by unvoiced consonants.

Vowels did not convey any information about grammatical category or number. The six vowels used to create the words were /u/, /i/, /o/, /a/, /e/, /E/.

Table 21.
Lexicon used in Experiment 4. Monosyllabic words were characterized by grammatical category and number. Words in parenthesis were used to study Lexical Generalization and Lexical and Syntactic Generalization.

<table>
<thead>
<tr>
<th>Grammatical Category</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nouns</td>
<td>du</td>
<td>ti</td>
</tr>
<tr>
<td></td>
<td>(ge)</td>
<td>(kE)</td>
</tr>
<tr>
<td>Verbs</td>
<td>bo</td>
<td>pa</td>
</tr>
</tbody>
</table>

Prosody: Words were then combined to build up sentences. Natural prosody was applied as in Experiment 3. A 25 ms pause was added between each word in order to facilitate segmentation and a 125 ms pause was added in order to mark the constituents’ boundaries.
Agreement Rule: The same hierarchical agreement rule as in Experiment 3 was applied (i.e., the verb (V) agrees with the highest noun (N) in its constituent; if there is no noun in its constituent, the verb agrees with the highest noun in the immediately preceding constituent). Examples of grammatical and ungrammatical sentences are given in Table 22.

The set of sentences used in this experiment included: 4 grammatical two-word sentences (NV); 4 ungrammatical two-word sentence; 8 grammatical three-word sentences (4 sentences N_NV and 4 sentences NN_V); 8 ungrammatical three-word sentences (4 sentences N_NV and 4 sentences NN_V); 8 grammatical four-word sentences (4 sentences NV_NV and 4 NVN_V); 8 ungrammatical four-word sentences (4 sentences NV_NV and 4 NVN_V). The sentences are listed in Appendix D.

Table 22.
Examples of sentences used in Experiment 4. The subject and verb for each sentence are written in bold font. Ungrammatical sentences are marked with the symbol *.

<table>
<thead>
<tr>
<th>Sentence structure</th>
<th>Grammatical sentences</th>
<th>Ungrammatical sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>NV</td>
<td>du bo</td>
<td>*du pa</td>
</tr>
<tr>
<td></td>
<td>kE pa</td>
<td>*kE bo</td>
</tr>
<tr>
<td>N_NV</td>
<td>du _ ti pa</td>
<td>*du _ ti bo</td>
</tr>
<tr>
<td></td>
<td>kE _ ge bo</td>
<td>*kE _ ge pa</td>
</tr>
<tr>
<td>NN_V</td>
<td>du ti _ bo</td>
<td>*du ti _ pa</td>
</tr>
<tr>
<td></td>
<td>kE ge _ pa</td>
<td>*kE ge _ bo</td>
</tr>
<tr>
<td>NV_NV</td>
<td>ti pa _ du bo</td>
<td>*ti pa _ du pa</td>
</tr>
<tr>
<td></td>
<td>ge bo _ kE pa</td>
<td>*ge bo _ kE bo</td>
</tr>
<tr>
<td>NVN_V</td>
<td>ti pa du _ pa</td>
<td>*ti pa du _ bo</td>
</tr>
<tr>
<td></td>
<td>ge bo kE _ bo</td>
<td>*ge bo kE _ pa</td>
</tr>
</tbody>
</table>

Locality: The hierarchical agreement rule led to different types of dependencies, namely adjacent dependency and non-adjacent dependency, as in Experiment 3.

Type of knowledge: In the present experiment we investigated Memorization, Lexical Generalization, Syntactic Generalization and Lexical and Syntactic Generalization.
Procedure

The experiment consisted of several sessions as schematized in Figure 13. In the Lexicon Familiarization Session, participants were familiarized with the lexicon; in the Rule Learning Session I, they were asked to listen to a sample of sentences following the hierarchical agreement rule and to perform a grammaticality judgment task. In the New Lexicon Familiarization Session, participants were familiarized with two new nouns of the lexicon; in the Rule learning Session II, they were asked to generalize the agreement rule to sentences constituted by two, three and four words.

Lexicon Familiarization Session

Participants were presented with trials introducing the words. Each trial had the following structure: a fixation cross was shown in the centre of the screen for 500 ms. Subsequently, the
word was presented auditorily followed by a 700 ms blank screen interval, which separated the trials. All words were presented once. Participants were asked to listen to the words carefully.

*Rule Learning Session I*

This session was constituted by two series of training and test phases repeated 3 times each, followed by a last test phase without training. The first series introduced two–word sentences (6 sentences in the training and 12 sentences in the test); the second series introduced three–word sentences (12 sentences in the training and 24 sentences in the test); the last test phase was constituted by four-word sentences (24 sentences).

**Training phases:** In this phase each trial had the following structure: a fixation cross was shown in the centre of the screen for 500 ms. A sentence was then presented auditorily. A 1000 ms blank screen interval separated the trials. Participants were asked to listen to the sentences carefully.

**Test phases:** In this phase each trial had the following structure: a fixation cross was shown in the centre of the screen for 500 ms. A sentence was presented auditorily and the response of the participant was recorded. If participants thought that the sentence was grammatical following what they had heard in the training phase, they had to press the green button; otherwise the red button. A feedback (✓ for correct answers and X for incorrect answers) lasting 1000 ms was then presented and a 500 ms blank screen interval separated the trials. Fifty percent of the sentences were grammatical.

*New Lexicon Familiarization*

This session was constituted by the same two phases as the Lexicon Familiarization Session, namely lexicon training and lexicon test. Two novel nouns were presented (i.e., ge, kE).

*Rule Learning Session II*

This session was constituted by one training phase and two test phases. Two words sentences constituted by the new nouns were presented in the training phase (18 sentences).
The two test phases were respectively constituted by three- and four-word sentences of the new lexicon (24 sentences each phase). Trials in the training and test phases had the same structure as those in Rule Learning Session I.

### 3.3.2. Results

**D-prime Analysis**

Before performing the statistical analyses we eliminated the outliers. A response was classified as an outlier if the reaction time was two or more SD from the subject’s mean for each sentence length (two-words; three-word; four-word). As a result of this procedure, 5.2% of the responses were dropped.

**Performance with respect to chance level**

Mean d-prime values for each type of knowledge and for each sentence structure are summarized in Table 23.

#### Table 23.

Mean d-prime values (SD in parenthesis) for each type of knowledge and sentence structure.

<table>
<thead>
<tr>
<th>Type of Knowledge</th>
<th>NV</th>
<th>N_NV</th>
<th>NN_V</th>
<th>NV_NV</th>
<th>NVN_V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memorization</td>
<td>2.23</td>
<td>0.52</td>
<td>0.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.44)</td>
<td>(0.78)</td>
<td>(0.71)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syntactic Generalization</td>
<td></td>
<td></td>
<td></td>
<td>0.78</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(2.00)</td>
<td>(1.11)</td>
</tr>
<tr>
<td>Lexical Generalization</td>
<td>0.98</td>
<td>-0.36</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.77)</td>
<td>(1.23)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexical and Syntactic Generalization</td>
<td>1.07</td>
<td>-0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.55)</td>
<td>(1.59)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In order to assess performance we performed several one-sample t-tests comparing the mean d-prime values to the chance level (i.e., as previously mentioned, zero is the chance level for d-prime values)\(^{27}\).

**Memorization** – Performance was significantly above chance for the NV structure \((t(29) = 8.481, p < .001)\); for the N_NV structure \((t(29) = 3.642, p < .001)\) and for the NN_V structure \((t(29) = 1.678, p = .052)\).

**Syntactic Generalization** – Performance was significantly above chance for the NV_NV structure \((t(29) = 2.769, p = .005)\) and marginally above chance for the NVN_V structure \((t(29) = 1.479, p = .075)\).

**Lexical Generalization** – Performance was significantly above chance for the N_NV structure \((t(29) = 3.052, p = .002)\). Performance for the NN_V structure was below chance \((t(29) = -1.606, p = .060)\).

**Lexical and Syntactic Generalization** – Performance was significantly above chance for the NV_NV structure \((t(29) = 2.939, p = .003)\) but not for the NVN_V structure \((t(29) = -0.144, p = .443)\).

**Performance as a function of manipulated variables**

We performed an ANOVA on d-prime values in order to study the Type of knowledge and Locality effects. There was a significant main effect of Locality, \(F(1,29) = 13.380, p = 0.001\). This effect reflected the fact that the performance of the participants was significantly better for the sentences with the adjacent dependency (i.e., N_NV and NV_NV), \((M = 0.84, SD = 1.6)\) than for the sentences with the non-adjacent dependency (i.e., NN_V and NVN_V), \((M = 0.06, SD = 1.2)\). The main effect of the Type of knowledge \((F < 1)\) and the interaction between Type of knowledge and Locality \((F(3,87) = 2.097, p = .106)\) were not significant.

\(^{27}\) The use of t-tests is justified by the fact that the distribution of d-prime values meets the assumption of normality \((p > .05\) for Kolmogorov-Smirnov tests) for each type of structure.
Template Analysis

The templates considered for this experiment were the same as the ones considered for Experiment 3.

Three-Word Sentences

As shown in Figure 14, two participants had the highest d-prime value for the hierarchical subject-verb agreement rule while the performance of the rest of the participants was not above chance for any of the templates.

![Figure 14](image.png)

**Figure 14.** Individual participant performance for the three-word sentences as a function of the template with the highest d-prime value. The line represents the threshold above which a performance was considered as above chance.

Four-Word Sentences

As shown in Figure 15, three participants had the highest d-prime value for the hierarchical subject-verb agreement rule while the performance of the rest of the participants was not above chance for any of the templates.
Chapter 3. The Hierarchical Syntactic Dependencies

Figure 15. Individual participant performance for four-words sentences as a function of the template with the highest d-prime value. The line represents the threshold above which a performance was considered as above chance.

Reaction Times Analysis

Mean RTs for correct responses each type of knowledge and sentence structure are summarized in Table 24.

Table 24. Mean RTs (SD in parenthesis) in ms for correct responses for each type of knowledge and sentence structure.

<table>
<thead>
<tr>
<th>Type of Knowledge</th>
<th>NV</th>
<th>N_NV</th>
<th>NN_V</th>
<th>NV_NV</th>
<th>NVN_V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memorization</td>
<td>607 (257)</td>
<td>784 (310)</td>
<td>785 (366)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syntactic Generalization</td>
<td></td>
<td>772 (396)</td>
<td>727 (410)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexical Generalization</td>
<td>762 (328)</td>
<td>790 (369)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexical and Syntactic</td>
<td>767 (294)</td>
<td>757 (357)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generalization</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
We performed an ANOVA on reaction times in order to study the within subject variables: Type of knowledge and Locality. There were no significant effects ($F_s < 1$).

### 3.3.3. Discussion

The main purpose of the present experiment was to study adults’ ability to learn a hierarchical agreement rule in the absence of a confounding condition such as the matching condition (i.e., intervening element with the same number as the subject), with a simplified lexicon and pauses marking words’ boundaries. In order to better understand the current results it is of interest to direct the attention to the different types of knowledge investigated in this experiment. For this reason, we centred our discussion on the results of the t-tests comparing participants’ performances to chance level.

In what concerns *Memorization*, we can observe that the performance on all three sentence structures (i.e., NV, N_NV, NN_V) was significantly above chance. These data refer to the participants’ capacity to judge sentence grammaticality on the basis of what they heard in the training phases. Participants might have judged the grammatical sentences as correct only because they recognized them from the training phase and the ungrammatical sentences as incorrect because they were novel. Therefore, they might have based their judgment simply on memorization, that is to say, on item-based knowledge. In order to move further in our understanding of the results we have to look at the other types of knowledge tested.

We observe that for *Syntactic Generalization* performance, participants were better in judging the sentence structure characterized by the adjacent dependency (i.e., NV_NV) than the one characterized by the non-adjacent dependency (i.e., NVN_N). As already mentioned in the discussion of the previous experiment, the NV_NV sentence structure is easier than the other structures, as it can be processed as two independent NV sentences. However, and crucially, performance was above chance for both structures. It is impossible to correctly judge these sentences on the basis of the linear order of the elements, since the grammatical order for one sentence structure (e.g., ti pa _ du bo) is ungrammatical for the other one (e.g., *ti pa du _ bo). Hence, these observations suggest that participants have learned the rule, that is, they must have integrated both the association between morphological markers and the fact
that their association depends on the position of the words in the constituent structure of the sentence, as a function of its prosody.

We consider now the data from the *Lexical Generalization* and from the *Lexical and Syntactic Generalization*. We observe that after having been trained on NV sentences with the two new nouns, participants performed well for the adjacent dependencies, both on three-word sentences (i.e., N_NV) and on four-word sentences (i.e., NV_NV). However, they failed in the judgment of the non-adjacent dependencies in three-word sentences and four-word sentences (i.e., NN_V and NVN_V). It is possible that the fact that participants have been trained only with NV sentences created a strict association between the two pairs of words (ge - bo and kE - pa), and that therefore participants judged three- and four-word sentences on the basis of the presence (or absence) of these pairs. This would account for the finding that in the NN_V sentences participants were significantly below chance level showing a systematic tendency for considering as correct sentences that contained one of the pairs, and as incorrect those that failed to contain one. This contrasts with the initial phase of the experiment during which participants were trained with NV as well as three-word sentences: their ability to extend the rule to new, four-word structures with that material may have relied on the presence of three-word structures during the training, which reduced the restricted pairwise association between the nouns and the verbs and pointed to the existence of non-adjacent dependencies.

To sum up, the results of the Syntactic Generalization condition suggest that participants were able to learn the rule and extend it to certain new structures. In contrast, the results of the Lexical Generalization and Lexical and Syntactic Generalization conditions indicate that they failed to generalize the rule to non-adjacent dependencies with a novel lexicon. These results could be explained considering the differences in the amount and in the type of dependencies presented in the training with the words du, bo, ti, pa, and in the training with the novel nouns ge, kE. For the novel nouns, participants were trained only with NV sentences. This might have influenced their ability to extend the rule to three- and four-word sentences with non-adjacent dependencies, as the training on the association between nouns and verbs was strictly reduced to the adjacent dependency.
We decided to make two major changes in the following experiment in order to improve the participants’ capacity to generalize the rule. In fact, even though participants learned the hierarchical agreement rule, they were still far from showing generalization to a new lexicon. The changes we will introduce aimed to improve participants’ mastery of the rule. The first modification relates to the type of markers and the nature of the association on which the morphological agreement was based. In experiments 3 and 4 the association between morphological agreement markers was established on the co-occurrence of syllables (i.e., the singular verb ending in –bo agrees with a singular noun ending in –du). In the following experiment the morphological agreement will be based on the identity relationship (i.e., repetition) between final vowels cueing number (i.e., the singular verb ending in –u agrees with a singular noun ending in –u). The second change concerns the choice of the phonological cue to mark grammatical categories distinction. We decided to employ syllable number to cue grammatical categories instead of the place of articulation of the consonants as the former is the way this information is conveyed in a few natural languages such as English. The above mentioned changes will be detailed in the next section.

3.4. Comparisons between Experiment 3 and Experiment 4

3.4.1. Results

D-prime Analysis

Mean d-prime values for each type of knowledge and sentence structure common to Experiment 3 and Experiment 4 are represented in Figure 16\textsuperscript{28}

\textsuperscript{28} For the experiment 3 we only considered the results in the mismatching condition.
We performed a mixed ANOVA on the d-prime values; the between-subject variable was the Experiment (i.e., Experiment 3, Experiment 4); the two within-subject variables were Type of knowledge (i.e., Lexical Generalization, Lexical and Syntactic Generalization) and Locality. The analysis revealed a main effect of Locality: $F(1,48) = 6.934, p = .011$. Performance was better for sentences with an adjacent dependency ($M = 0.67, SD = 1.7$) than for sentences with a non-adjacent dependency ($M = 0.03, SD = 1.17$). The main effects of Type of knowledge ($F(1,48) = 2.232, p = .142$) and Experiment ($F < 1$) were not significant. There was a significant interaction between Locality and Experiment: $F(1,48) = 6.029, p = .018$. The difference between Experiments 3 and 4 was significant for the adjacent dependency ($t(19) = 2.207, p = .040$), while there was only a tendency for non-adjacent dependencies ($t(19) = 1.889, p = .074$). The interaction between Type of knowledge and Locality ($F < 1$), Type of knowledge and Experiment ($F < 1$) and between Type of knowledge, Locality and Experiment ($F(1,48) = 2.165, p = .148$) were not significant.

**Reaction Time Analysis**

We performed a mixed ANOVA on the reaction times; the between-subject variable was the Experiment (i.e., Experiment 3, Experiment 4); the two within-subject variables were Type of knowledge (i.e., Lexical Generalization, Lexical and Syntactic Generalization) and
Locali

ty. The analysis revealed a significant main effect of Type of knowledge: $F(1,48) = 7.778, p = .008$. Participants were faster in judging sentences in the Lexical and Syntactic Generalization ($M = 696, SD = 369$) than sentences in the Lexical Generalization ($M = 796, SD = 346$). The main effects of Locality ($F(1,48) = 1.563, p = .217$) and Experiment ($F < 1$) were not significant. There was a significant interaction between Type of knowledge and Experiment: $F(1,48) = 5.738, p = .021$. The difference between Lexical Generalization and Lexical and Syntactic Generalization was significant for Experiment 3 ($t(19) = 5.085, p < .001$), but not for Experiment 4 ($t < 1, ns$). The interaction between Locality and Experiment ($t < 1, ns$), the interaction between Type of knowledge and Locality ($t < 1, ns$), and the interaction between Type of knowledge, Locality and Experiment ($F(1,48) = 1.346, p = .252$) were not significant.

3.4.2. Discussion

The ANOVAs presented above allow only a partial comparison between Experiment 3 and Experiment 4 as they took into account only the Lexical Generalization and the Lexical and Syntactic Generalization types of knowledge. Since we observed that in Experiment 4 participants were successful in the Syntactic Generalization condition, which is not considered in the present comparison, we have to keep in mind that we compared Experiment 3 with the weakest results in Experiment 4. Nevertheless, we observed an improvement in the performance to the adjacent dependencies in Experiment 4 compared with Experiment 3. This observation is in line with the idea that the training with the NV sentences with the novel lexicon in Experiment 4 facilitated the pairwise association between the nouns and the verbs and consequently contributed to judgment of the sentences characterized by adjacent dependencies. The assumption that the strength of this association also prevented participants from learning the non-adjacent dependencies is supported by the fact that there was a tendency for non-adjacent dependencies to be judged as worse in Experiment 4 than in Experiment 3.
3.5. Experiment 5

The results from Experiment 4 suggested that participants were able to learn the hierarchical agreement rule, but failed to generalize this rule to a new lexicon and syntactic structure (Lexical Generalization and Lexical and Syntactic Generalization). Since we assumed that the identification of grammatical categories and number markers, and the learning of the morphological agreement markers relationship could be improved compared to the previous experiment, we decided to convey this information in a more salient way. Consequently, in order to ameliorate the identification of these characteristics and improve rule learning we decided once again to change some aspects of the materials. In particular, we changed two main characteristics of the materials compared to Experiment 4. The first modification concerned the use of final vowels to mark number and the second modification concerned the use of syllable number to mark the distinction between grammatical categories.

In the present experiment number was marked by final vowels (i.e., -u for singular words and -i for plural words) and not by the voicing feature of the consonants as in the previous experiment. Consequently, if the morphological agreement markers relationship in Experiment 4 was based on the consonants (or on the syllables constituted by a consonant and a vowel), the morphological agreement markers relationship in Experiment 5 involved vowels. In particular, while in Experiments 3 and 4 the relation between morphological agreement markers was based on a simple association (i.e., co-occurrence) between the markers, in the present experiment it was based on identity. Therefore, in Experiment 5 the identity between morphological agreement markers was implemented as the identity of the final vowels (i.e., the same vowel for singular nouns and verbs, and the same vowel for plural

\[29\] We chose to use vowels over consonants because it has been shown that consonants and vowels play a different role in language acquisition, with the latter carrying grammatical information (Bonatti et al., 2005; Toro et al., 2008).

\[30\] We already used this type of relationship when we investigated the learning of the linear agreement rule in Experiment 2. The participants to that experiment did not successfully learn the rule and we assumed that learning was prevented by the presence of a perceptual grouping bias (i.e., they applied the rule: all the words must have the same final vowel). Two factors might have induced this bias, the presence of the matching condition and the way we marked number (i.e., same final vowel for singular words, same final vowels for plural words). As we thought that the main problem lied in the latter we decided to change in the following experiments the number markers and consequently the agreement markers relationship. In light of the results of Experiment 3 which suggested that the match condition prevented learning, we think that it is reasonable to use again this type of relationship between agreement markers.
nouns and verbs). We assumed that the identity relationship was stronger than the association relationship. In fact, we can define association as the link between elements in memory, while identity is a special case of association in which elements are linked in a perceptual and salient way, namely by the condition of being the same.

As a consequence of the identity relationship between final vowels, agreement was based on the repetition of the same marker. Interestingly, it has been shown that repetition plays a facilitatory role in language acquisition (see for example Endress et al., 2007; Gervain, Macagno, et al., 2008; Mehler et al., 2008; Monaghan & Rowson, 2008 among others). Moreover, very often the first words learned by children are characterized by reduplications (e.g., bébé (baby) and dodo (sleep) in French or papà (dad) and mamma (mum) in Italian). Also, several studies testing adults in artificial grammar learning paradigms demonstrated that the presence of repetitions influences the capacity to generalize a finite state grammar to a novel lexicon (Endress et al., 2009; Lotz & Kinder, 2006). If these patterns are not available, the generalization to a new lexicon is not possible (Gomez et al., 2000; Tunney & Altmann, 2001). Therefore, repetitions seem to have a particular status that facilitates the extraction and generalization of syntactic dependencies. By implementing number features across nouns and verbs in terms of morphological identity rather than in terms of association between syllables, we expected to reduce the cognitive burden involved in the learning of morphological agreement. We also assumed that this would result in improved rule learning, since participants’ resources will all be allocated to the last learning step that learners have to perform. As mentioned in the introduction of Chapter 3, in this step participants have to parse the sentence into constituents, identify the syntactic roles, and learn the agreement between the subject and the verb.

The second change in the materials of the present experiment concerns the phonological cue used to implement the distinction between grammatical categories. In Experiments 3 and 4 this information was carried by the features of consonants and vowels on which the syntactic dependency was determined. In Experiment 5 we used a difference in syllable number to distinguish nouns and verbs. Interestingly, previous studies (Cassidy & Kelly, 1991; Kelly, 1992, 1996; Monaghan et al., 2005; Shi et al., 1998) showed that phonological correlates to grammatical categories can be found in natural language. For
instance, Kelly (1992, 1996) observed that in English the number of syllables in a word is a reliable cue to distinguish nouns and verbs; nouns tend to have more syllables than verbs. In another paper, Cassidy and Kelly (1991) examined the contribution of phonology to word categorization. In the first two studies they analyzed samples of parental speech to children and adult written prose. The analysis revealed that polysyllabic nouns were common while polysyllabic verbs were not. In two other studies, they investigated adults and children’s sensitivity to this phonological cue. They tested the participants in a paradigm in which they had to categorize nonsense words. The results showed that, when the word was monosyllabic, it was more often associated to the verb category rather than when it was trisyllabic. This phonological correlate might work as cue to categorization for infants acquiring their own language. In our experiments, in order to perform proficiently with respect to the hierarchical agreement rule learning, the correct assignment of the grammatical category to the words is necessary. Therefore, in the present experiment we decided to facilitate this task using the same phonological cue as natural language to distinguish grammatical categories; nouns were bisyllabic words and verbs were monosyllabic words.

To sum up, we hypothesized that the set of novel cues introduced in the present experiment (i.e., final vowels identity and syllable number) would improve the participants’ ability to learn and generalize the hierarchical agreement rule. The main properties of Experiment 5 are summarized in Table 25. We expected that participants would be able to learn successfully the rule, that is their performance would be significantly above chance. Finally in the present experiment a video accompanied the presentation of words and sentences in the training phases. We added this visual stimulus in order to be able to compare this experiment with Experiment 6, in which we introduced videos carrying semantic information.
Table 25.  
Main properties of the materials used in Experiment 5.

<table>
<thead>
<tr>
<th>Stimuli Properties</th>
<th>Experiment 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vocabulary</strong></td>
<td></td>
</tr>
<tr>
<td>Bisyllabic: CVC-V (nouns)</td>
<td></td>
</tr>
<tr>
<td>Monosyllabic: CCV (verbs)</td>
<td></td>
</tr>
<tr>
<td><strong>Grammatical Category</strong></td>
<td>Number of syllables and syllable structure:</td>
</tr>
<tr>
<td><strong>Markers</strong></td>
<td>Nouns: CVC-V</td>
</tr>
<tr>
<td></td>
<td>Verbs: CCV</td>
</tr>
<tr>
<td><strong>Number</strong></td>
<td>Final vowels:</td>
</tr>
<tr>
<td><strong>Markers</strong></td>
<td>Singular: /u/</td>
</tr>
<tr>
<td></td>
<td>Plural: /i/</td>
</tr>
<tr>
<td><strong>Prosody</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Pitch</strong></td>
<td>Natural</td>
</tr>
<tr>
<td><strong>Pauses</strong></td>
<td>20ms between words</td>
</tr>
<tr>
<td></td>
<td>200ms between constituents</td>
</tr>
<tr>
<td><strong>Agreement</strong></td>
<td>Hierarchical</td>
</tr>
<tr>
<td><strong>Rule</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Locality</strong></td>
<td>Adjacent and Non-adjacent</td>
</tr>
<tr>
<td><strong>Match</strong></td>
<td>Mismatching</td>
</tr>
<tr>
<td><strong>Type of Knowledge</strong></td>
<td>Memorization</td>
</tr>
<tr>
<td><strong>Lexical Generalization</strong></td>
<td>Lexical and Syntactic Generalization</td>
</tr>
</tbody>
</table>

3.5.1. Method

Participants

Twenty participants took part in the experiment. They were all monolingual French speakers, with ages ranging between 18 and 29, with no reported hearing or language impairment. They were paid or they received course credit for their participation.

Materials

Vocabulary: A lexicon of 20 bisyllabic and monosyllabic words was created. All words were characterized by grammatical category (noun or verb) and number (singular or plural). The lexicon was constituted by four singular nouns (i.e., meku, Rabu, govu, lEpu), and four plural nouns (i.e., meki, Rabi, govi, lEpi), two singular verbs (i.e., dRu, stu) and two plural verbs (i.e., dRi, sti). Eight additional nouns (i.e., meka, meko, Raba, Rabo, gova, govo, lEpa, lEpo) were created for the lexicon test and for the new lexicon test. The words used in the rule
learning sessions are listed in Table 26. Grammatical category and number information were marked by phonological cues as detailed below.

**Grammatical category marker:** The distinction between grammatical categories was cued by the *number of syllables* and *syllable structure*, namely bisyllabic words (CV-CV) for nouns and monosyllabic words (CCV) for verbs.

**Number marker:** The distinction between singular and plural words was implemented using different *final vowels*, namely /u/ for singular words and /i/ for plural words.

**Table 26.**

*Words used in Experiment 5. The number marker is highlighted in bold font. Words in parenthesis were used to study Lexical Generalization and Lexical and Syntactic Generalization.*

<table>
<thead>
<tr>
<th>Grammatical Category</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nouns</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>me-ku</td>
<td>me-ki</td>
<td></td>
</tr>
<tr>
<td>Ra-bu</td>
<td>Ra-bi</td>
<td></td>
</tr>
<tr>
<td>(go-vu)</td>
<td>(go-vi)</td>
<td></td>
</tr>
<tr>
<td>(LE-pu)</td>
<td>(LE-pi)</td>
<td></td>
</tr>
<tr>
<td><strong>Verbs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dRu</td>
<td>dRi</td>
<td></td>
</tr>
<tr>
<td>(stu)</td>
<td>(sti)</td>
<td></td>
</tr>
</tbody>
</table>

**Prosody:** Words were combined to build up sentences. In order to obtain a natural prosody to apply to these artificial sentences we used the same methodology as in Experiments 3 and 4. The only difference was that the pitch contour of the sentence structures N_NV and NV_NV was adjusted in order to have the same contour as that for the NV sentences. This was possible by reusing the values of the pitch contour of the NV recordings with MBROLA. We decided to do this to keep the prosody constant as much as possible for the same structures. A 20 ms *pause* was added between each word in order to facilitate lexical segmentation and a longer pause (200 ms) was added in order to mark the constituents’ boundaries.

**Agreement Rule:** The same hierarchical agreement rule as in experiments 3 and 4 was applied (i.e., the verb (V) agrees with the highest noun (N) in its constituent; if there is no noun in its
constituent, the verb agrees with the highest noun in the immediately preceding constituent). Examples of grammatical and ungrammatical sentences are given in Table 27.

The complete set of sentences included: 8 grammatical two-word sentences (N,Vi); 8 ungrammatical two-word sentences (Ni,Vj); 16 grammatical three-word sentences (8 sentences Ni,Ni,Vj and 8 sentences Nj,Nj,Vj); 16 ungrammatical three-word sentences (8 sentences Ni,Ni,Vj and 8 sentences Nj,Nj,Vj); 8 grammatical four-word sentences (4 sentences Ni,Vj,Nj,Vj and 4 Nj,Vj,Nj,Vj); 8 ungrammatical four-word sentences (4 sentences Ni,Vj,Nj,Vj and 4 Nj,Vj,Nj,Vj). The sentences are listed in Appendix E.

Table 27.
Examples of sentences used in Experiment 5. The subject and verb for each sentence are written in bold font. Ungrammatical sentences are marked with the symbol *.

<table>
<thead>
<tr>
<th>Sentence structure</th>
<th>Grammatical sentences</th>
<th>Ungrammatical sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>NV</td>
<td>govu stu Rabi dRi</td>
<td>*govu sti *Rabi dRu</td>
</tr>
<tr>
<td>N_NV</td>
<td>Rabi _ meku dRi govu _ lEpi sti</td>
<td>*Rabi _ meku dRi *govu _ lEpi stu</td>
</tr>
<tr>
<td>NN_V</td>
<td>Rabi meku _ dRi govu _ lEpi _ stu</td>
<td>*Rabi meku _ dRi *govu _ lEpi _ sti</td>
</tr>
<tr>
<td>NV_NV</td>
<td>lEpu _ govi _ lEpu _ sti</td>
<td>*lEpu _ govi _ lEpu _ sti</td>
</tr>
<tr>
<td>NVN_V</td>
<td>lEpu _ govi _ lEpu _ sti</td>
<td>*lEpu _ govi _ lEpu _ sti</td>
</tr>
</tbody>
</table>

Locality: The hierarchical agreement rule led to different types of dependencies, namely adjacent dependency and non-adjacent dependency, as in Experiments 3 and 4.

Type of knowledge: In the present experiment we investigated Memorization, Lexical Generalization, and Lexical and Syntactic Generalization.

Visual Stimuli: A video was created for this experiment. It was used in all the training phases. The video lasted 3600 ms and showed a square moving in a random trajectory. The video was produced with Apple Keynote and then converted to an .avi file with Windows Movie Maker.
Procedure

The experiment consisted of several sessions each one made up of training and test phases as schematized in Figure 17. This experiment was made up of four different sessions. In the Lexicon Learning Session participants were familiarized with four nouns of the lexicon; in the Rule Learning Session I, they were asked to listen to a sample of sentences following the hierarchical agreement rule and to perform a grammaticality judgment task; in the New Lexicon Learning Session participants were familiarized with four new nouns of the lexicon; in the Rule Learning Session II they were asked to generalize the agreement rule for sentences constituted by two, three and four words.

Figure 17. Schematic representation of Experiments 5.
Lexicon Learning Session

Lexicon training phase: Participants were presented with 24 trials introducing four singular and plural nouns (i.e., meku, meki, Rabu, Rabi). Each trial had the following structure: a fixation cross was shown in the centre of the screen for 500 ms, subsequently, the video was presented followed by a 500 ms blank screen interval which separated the trials. The videos in this phase were characterized by the following features: they lasted 3600 ms; the square was on the screen and it started moving, after 640 ms one word was presented auditorily. Trials were presented in a randomized order. Participants were asked to listen to the words and watch the videos carefully.

Lexicon test phase: Participants were asked to perform a word 2-choice recall task. They had to indicate, by pressing buttons on a response box, whether the word they were listening to had already been presented during the lexicon training phase, in which case they had to press the button labelled “O”. If it hadn’t, they had to press the button labelled “N”. They were presented with 24 trials. Each trial had the following structure: a fixation cross was shown in the centre of the screen for 500 ms, subsequently a word was presented auditorily, and the answer was recorded.

Rule Learning Session I

This session was constituted by two series of rule training and rule test phases, the first series introduced sentences with two words (12 sentences in the training and 24 sentences in the test) and the second series introduced sentences with three words (24 sentences in the training and 48 sentences in the test).

Training phases: In this phase, each sentence was presented twice: once before the video and once during the video. More precisely, each trial had the following structure: a fixation cross was shown in the centre of the screen for 500 ms. A sentence was then presented auditorily followed by a 500 ms blank screen interval and the video. Videos lasted 3600 ms. The square was already on the screen at the beginning of the video. The sentence presentation started 720 ms after the beginning of the video. A 1000 ms blank screen interval separated the trials. All
sentences presented in this phase were grammatical. Participants were asked to listen to the sentences and watch the videos carefully.

**Test phases:** In this phase each trial had the following structure: a fixation cross was shown in the centre of the screen for 500 ms. Subsequently, the sentence was presented auditorily followed by a question mark which stayed on the screen until a response was given, or for a maximum of 5000 ms. If the participants thought that the sentence was grammatical according to what they had heard in the training phase, they had to press the green button (otherwise the red button) on the response box. A feedback (✓ for correct answers and X for incorrect answers) lasting 1000 ms was then presented and a 1000 ms blank screen interval separated the trials. Fifty percent of the sentences were grammatical.

*New Lexicon Learning Session*

This session was constituted by the same two phases as the Lexicon Learning Session, namely lexicon training and lexicon test. Four novel nouns were introduced (i.e., govu, govi, lEpu, lEpi).

*Rule Learning Session II*

This session comprised a series of rule training and test phases in which participants were presented and tested with two-word sentences (12 sentences in the training and 24 sentences in the test) containing the novel nouns introduced in the New Lexicon Learning Session and a novel verb. Subsequently, participants were directly tested for three-word sentences (48 sentences) and for four-word sentences (48 sentences). Fifty percent of the sentences in the rule test phases were grammatical.

### 3.5.2. Results

*D-prime Analysis*

Before performing the statistical analyses, we eliminated the outliers. A response was defined as an outlier if the reaction time was two or more SD from the subject’s mean for each
sentence length (2 words; 3 words; 4 words). As a result of this procedure, 5.1% of the responses were dropped.

**Lexicon Learning**

All participants performed well on the lexicon learning tests. The mean percentage of correct responses to the word recall task was 95% ($SD = 6.4$) for the first lexicon test and 96% ($SD = 5.8$) for the new lexicon test. The performance did not differ significantly between the first and the second lexicon test ($z = 0.528$, $N – Ties = 12$, $p = .597$, two-tailed).

**Performance with respect to chance level**

Mean d-prime values for each type of knowledge and structure are summarized in Table 28.

<table>
<thead>
<tr>
<th>Type of Knowledge</th>
<th>NV</th>
<th>N_NV</th>
<th>NN_V</th>
<th>NV_NV</th>
<th>NVN_V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memorization</td>
<td>5.09</td>
<td>1.54</td>
<td>1.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.03)</td>
<td>(1.83)</td>
<td>(1.75)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexical Generalization</td>
<td>1.21</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.74)</td>
<td>(2.44)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexical and Syntactic Generalization</td>
<td>1.70</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.68)</td>
<td>(1.42)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In order to test the rule learning we performed several one-sample t-tests\(^{31}\) comparing the mean d-prime values to the chance level (i.e., as previously mentioned, zero is the chance level for d-prime values).

**Memorization** – Performance was significantly more accurate than chance for the NV structure ($t(19) = 22.103$, $p < .001$), for the N_NV structure ($t(19) = 3.752$, $p < .001$), and for the NN_V structure ($t(19) = 2.680$, $p = .007$).

\(^{31}\) The use of t-tests is justified by the fact that the distribution of d-prime values meets the assumption of normality ($p > .05$ for Kolmogorov-Smirnov tests) for each type of structure.
**Lexical Generalization** – Performance was significantly more accurate than chance for the N_NV structure ($t(19) = 3.119, p = .003$) and marginally for the NN_V structure ($t(19) = 1.533, p = .071$).

**Lexical and Syntactic Generalization** – Performance was significantly more accurate than chance for the NV_NV structure ($t(19) = 4.524, p < .001$), but not for the NVN_V structure ($t(19) = 0.213, p = .416$).

**Performance as a function of the manipulated variables**

We performed an ANOVA on d-prime values in order to study the type of knowledge and locality effects. There was a main effect of locality, $F(1,19) = 11.577, p = .003$. Performance was better for sentences with an adjacent dependency ($M = 1.48, SD = 1.73$) than for sentences with a non-adjacent dependency ($M = 0.65, SD = 1.93$). The main effect of type of knowledge ($F < 1$) and the interaction between the variables ($F(2,38) = 2.286, p = .116$) were not significant.

**Template Analysis**

The templates considered for this experiment were the same as the ones considered for Experiment 3.

**Three-Word Sentences**

As shown in Figure 18, seven participants had the highest d-prime value for the hierarchical subject-verb agreement rule and they performed above chance. One participant performed above chance with the template $2^\wedge-3^\wedge$. The performance of twelve participants was not above chance for any of the templates.
Figure 18. Individual participant performance for the three-word sentences as a function of the template with the highest d-prime value. The line represents the threshold above which a performance was considered above chance.

Four–Word Sentences

As shown in Figure 19, five participants had the highest d-prime value for the hierarchical subject-verb agreement rule and they performed above chance. Six participants performed above chance with the template $3^\wedge-4^\wedge$. The performance of twelve participants was not above chance for any of the templates.
Figure 19. Individual participant performance for four-word sentences as a function of the template with the highest d-prime value. The line represents the threshold above which a performance was considered above chance.

Reaction Times Analysis

Mean RTs for correct responses for each type of knowledge and structure are summarized in Table 29.

Table 29.
Mean RTs in ms (SD in parenthesis) for each type of knowledge and sentence structure.

<table>
<thead>
<tr>
<th>Type of Knowledge</th>
<th>NV</th>
<th>N_NV</th>
<th>NN_V</th>
<th>NV_NV</th>
<th>NVN_V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memorization</td>
<td>477 (190)</td>
<td>1006 (490)</td>
<td>1066 (472)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexical Generalization</td>
<td>915 (403)</td>
<td>953 (358)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexical and Syntactic</td>
<td>938 (381)</td>
<td>919 (452)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
We performed an ANOVA on reaction times in order to study type of knowledge and locality effects. None of the effects resulted significant ($F_s < 1$).

### 3.5.3. Discussion

The aim of the present experiment was to investigate the capacity of adults to learn a hierarchical agreement rule in an artificial language. In particular, we examined the learning under specific circumstances, namely when the morphological agreement markers relationship was based on the identity between final vowels and grammatical category was cued by the number of syllables constituting the words. We will first focus on the results of the t-tests comparing participants performance to chance level for each type of knowledge tested, namely Memorization, Lexical Generalization, and Lexical and Syntactic Generalization and the template analysis.

With respect to **Memorization**, results indicated that participants performed above chance for all the structures investigated, namely NV, N_NV, NN_V. For the **Lexical Generalization**, participants performed above chance for the N_NV sentences and participants also showed a tendency to perform above chance for the NN_V sentences. Finally, for the **Lexical and Syntactic Generalization**, we found the same pattern of results as in the previous experiment. Performance in the grammaticality judgment task was above chance for the NV_NV sentences, but not for the NVN_V sentences. Once again, when the hierarchical rule led to an adjacent dependency, the sentences were judged better than when it led to a non-adjacent dependency. This result was highlighted by the main effect of locality in the analysis of variance. Considering these results, we can reasonably think that participants learned the rule and its generalization to a novel lexicon. However, they failed when they had to generalize the rule both to a novel lexicon and to new syntactic structures.

We consider now the pattern of results that emerged in the template analysis. The template analysis for the sentences constituted by four words suggested that a group of six participants applied a linear rule (i.e., the third and the fourth words have the same number) to perform the grammaticality judgment task. This observation may well explain the difference in the performance between NV_NV and NVN_V structures. The NV_NV sentences could be
judged correctly applying this linear rule. However, this was not the case for the NVN_V sentences. It is thus possible that these participants did not learn the hierarchical agreement rule, but applied the linear one.

The template analysis also showed that there was a group of seven participants who successfully learned the rule and managed to generalize that to a novel lexicon (see the template analysis for three-word sentences), and another smaller group of five participants who generalized the rule to novel lexicon and syntactic structures (see the template analysis for four-word sentences). The participants’ performance appears to be highly heterogeneous. However, we can reasonably propose to classify them according to different profiles. 1) One profile refers to participants who did not learn the rule or any other template. Participants belonging to this group did not reach the chance level for the rule or any of the proposed templates. 2) The second profile refers to participants who learned a rule different from the hierarchical agreement rule. These participants performed above chance for one of the alternative templates. 3) The third profile refers to participants who learned the rule, but managed to generalize it only to a novel lexicon. These participants are those performing above chance with the rule for the three-word sentences but not for the four-word sentences. 4) The fourth profile refers to participants who performed well only in the lexicon and syntactic generalization. We can reasonably think that these participants, although slower than others, learned the rule. 5) Finally, the last profile refers to the participants who learned and generalized the rule to a novel lexicon and new syntactic structures. These participants performed above chance for the rule both for the three-word sentences and the four-word sentences.

The fact that we can identify so many different profiles in the participants’ behaviour is tightly linked to the question about the nature of the knowledge acquired in ALL paradigms. The general idea is that, if participants successfully learn a rule, they should have a good performance in every condition of the experiment (i.e., all types of knowledge and sentence structures). If we observe different performance profiles, we should wonder about what is ultimately learned by those participants. Are they learning alternative rules? Are they

---

32 Since we do not exclude that other “templates” that we did not identify are possible, we cannot exclude that these participants were applying a different incorrect rule to perform the grammaticality judgment task.
memorizing fragments of the training items? Are they judging the similarity between the training and the test items? We will treat this topic in depth in the General Discussion chapter.

To sum up, the present results showed that there has been an improvement in the ability of adults to learn a hierarchical agreement rule and generalize that to a new lexicon. However, this learning is characterized by a poor performance for the Lexical and Syntactic Generalization. We observed a high variability in the group of participants that might have prevented the learning effect for the Lexical and Syntactic Generalization to emerge statistically. As a matter of fact, several different profiles characterizing participants’ performance emerged. Once again we have shown that learning a hierarchical agreement rule is difficult and adult learners need cues that highlight the relevant units involved in the syntactic dependency. Since performance is still far from being a 100% successful, we decided to investigate the facilitatory role of an additional cue. With Experiment 6, we introduced semantics and thereby aimed to test whether this cue allows participants to reach a higher level of command of the rule.
3.6. Comparisons between Experiment 4 and Experiment 5

3.6.1. Results

*D-prime Analysis*

Mean d-prime values for each type of knowledge and sentence structure common to Experiment 4 and Experiment 5 are represented in Figure 20.

![Figure 20](image)

**Figure 20.** Mean d-prime values for each type of knowledge and sentence structure common to Experiment 4 and Experiment 5 (Error bars represent half standard deviation).

We performed a mixed ANOVA on the d-prime values; the between-subject variable was the *Experiment* (i.e., Experiment 4, Experiment 5); the two within-subject variables were *Type of knowledge* (i.e., Memorization, Lexical Generalization, Lexical and Syntactic Generalization) and *Locality*. The analysis revealed that the main effect of Experiment was significant: $F(1,48) = 6.662, p = .013$. Performance was better in Experiment 5 ($M = 1.07, SD = 1.9$) than in Experiment 4 ($M = 0.4, SD = 1.43$). The main effect of the Locality was also significant: $F(1,48) = 24.468, p < .001$. Performance was better for sentences with an adjacent dependency ($M = 1.17, SD = 1.7$) than for sentences with a non-adjacent dependency ($M = 0.3, SD = 1.5$). The main effect of Type of knowledge was not significant ($F < 1$). There was a significant interaction between Type of knowledge and Locality: $F(2,96) = 3.129, p = .048$. The effect of Locality was significant for the three types of knowledge...
(Memorization: $t(49) = 2.117, p = .039$; Lexical Generalization: $t(49) = 2.976, p = .005$; Lexical and Syntactic Generalization: $t(49) = 4.150, p < .001$), but it was bigger for the Lexical and Syntactic Generalization (Adjacent-Non-adjacent difference = 1.31) than for the Lexical Generalization (Adjacent-Non-adjacent difference = 0.96) and the Memorization (Adjacent-Non-adjacent difference = 0.38). The interaction between Experiment and Type of knowledge was not significant ($F < 1$), as well as the interaction between Experiment and Locality ($F < 1$) and the interaction between Type of knowledge, Locality and Experiment ($F(2,96) = 2.035, p = .136$).

**Reaction Times Analysis**

We performed a mixed ANOVA on the reaction times; the between-subject variable was the *Experiment* (i.e., Experiment 4, Experiment 5); the two within-subject variables were *Type of knowledge* (i.e., Memorization, Lexical Generalization, Lexical and Syntactic Generalization) and *Locality*. The Greenhouse-Geisser Epsilon was used in order to deal with lack of sphericity for the factor Type of knowledge. The main effect of the variable Experiment was significant $F(1.78, 85.42) = 1.120, p = .326$. Participants were faster in Experiment 4 than in Experiment 5. None of the remaining main effects resulted significant: Type of Knowledge $F(1.78, 85.42) = 1.120, p = .326$, Locality $F < 1$.

**3.6.2. Discussion**

The present comparison confirmed the observations already presented in the discussion of Experiment 5. The main effect of the variable Experiment confirmed that performance in Experiment 5 was better than performance in Experiment 4. Once again the main effect of Locality highlighted the fact that adjacent dependencies are easier to learn than non-adjacent dependencies. Participants’ learning of the hierarchical agreement rule showed an improvement in Experiment 5 compared to Experiment 4, especially for what concerns the NN_V sentences. It is possible that the use of a different phonological cue to mark the grammatical category and the identity between the final vowels marking number agreement improved rule learning. There were also other minor changes that have been made in the present experiment compared to the previous one (i.e., longer pauses at constituents’ boundaries, common pitch contour for NV structures embedded in other sentences). As a
consequence, we cannot determine which variable(s) played the main role in this improvement with certainty. However, we can reasonably think that the novel cues, and especially the fact that the morphological agreement was determined by the identity relation between the final vowels, had a fundamental impact on the participants’ performance. The salience of the identity relation compared to the simple association used in the previous experiment might have played the largest role in learning.
Chapter 4. The Role of Semantic Cues
4.1. Introduction

On the basis of the previous experiments’ results, we concluded that, whereas learning and generalizing a hierarchical subject-verb agreement rule in the context of an artificial language is not an easy task, it seems possible when the appropriate set of cues is given. However, the performance to the grammaticality judgment task in the previous experiment is still far from being excellent (i.e., the mean performance was never at a 100% of correct responses). This observation raises questions about what is ultimately learned during ALL experiments, and which is the type of knowledge acquired by the participants. As Crain and Thornton (1998) pointed out, if participants acquire perfect knowledge of a rule, they should show only little deviation from 100% correct performance. However, this seems not be the case in our experiments. Therefore, since we want to investigate how to make this kind of rule learnable to a higher level of command than in the previous experiments, we decided to add an additional cue. In this experiment we decided to investigate the role of semantics.

One of the reasons why artificial languages are used in research is that they allow disentangling aspects of language such as syntax, prosody, and semantics. They allow controlling the learning conditions to which participants are exposed to. However, it is impossible to learn complex syntactic rules, as the hierarchical agreement one, without the aid of multiple cues and constraints (e.g., phonological, prosodic, etc.) in the artificial language input. For instance, it has been shown that, in order to learn the relationships between adjacent elements, statistical information was sufficient (Saffran, Aslin, et al., 1996; Saffran, Newport, et al., 1996). However, to learn relationships involving distant elements, additional perceptual cues were necessary to differentiate the elements involved in the dependency from the interveners (Creel et al., 2004; Endress, 2010; Gomez, 2002; Newport & Aslin, 2004; Peña et al., 2002).

In the previous experiments the relevance of different phonological cues to mark the category of the words and prosodic cues to mark the constituent structure of the sentences were explored. In this experiment we investigated the relevance of another important cue, which is semantics, to the learning of the hierarchical agreement rule in an artificial language
learning paradigm. The role of semantics in natural language acquisition has been extensively studied, and it has been proposed that semantics and syntax interact during language acquisition. For instance, the semantic bootstrapping hypothesis (Pinker, 1984, 1987; Wexler & Culicover, 1980) claims that children may use semantics to predict syntax. This hypothesis assumes that infants have an innate access to concepts like “object”, “person”, and “action”. It also assumes that there is a systematic correspondence between semantic concepts and the way they are expressed; for instance, words referring to objects are nouns and words referring to actions are verbs. Therefore, infants can infer the grammatical categories of the words from these concepts. Subsequently, they assign the thematic roles to the arguments in sentences. This basic package of knowledge allows them to deduce more complex syntactic notions that are not mapped transparently in the world. In sum, the semantic bootstrapping hypothesis proposes that children can exploit semantics concepts to deduce syntactic ones. Indeed, the reference world may represent a source of information which facilitates syntax learning; however it is still debated whether it represents a necessary source of information. Interestingly, as Gleitman (1990) proposed, it is possible that semantic bootstrapping contributes to the acquisition of syntax in a complementary mode with other procedures.

The role of semantics has not been extensively investigated in ALL paradigms. To our knowledge only a few studies (Moeser & Bregman, 1972; Morgan & Newport, 1981; Morgan et al., 1987; Mori & Moeser, 1983) analyzed the role of semantics in syntactic rule learning in an artificial language learning paradigm. In other studies, researchers, primarily interested in defining the brain correlates of language learning and syntactic processing, included semantics in their miniature languages. They assumed that semantics would have facilitated learning (see for instance Amato & MacDonald, 2010; Friederici, Rüschmeyer, Hahne, & Fiebach, 2003; Friederici, Steinhauer, & Pfeifer, 2002; Opitz & Friederici, 2004), although they did not directly investigate its role.

Two studies (Moeser & Bregman, 1972; Morgan & Newport, 1981) are particularly interesting and allow us to introduce our experiment. Moeser and Bregman (1972) tested adults’ capacity to learn a phrase structure grammar in different conditions which varied in the
presence of semantic references. They used visual features (e.g., shapes, lines, orientation) to mark reference of nonsense words and tested the learning of the artificial grammar in three conditions. Their results showed that a complex grammar was learned only when there was a correlation between syntax and semantics, that is when the syntactic dependencies between words were marked by the combination of visual features. Accuracy was around 79% when there was a syntax-semantics correlation, while it was around 47% when there was an arbitrary correspondence between words and visual features. The authors concluded that the learning of a complex grammar was possible only when syntactic dependencies were semantically represented.

On the other hand, Morgan and Newport (1981) argued that it was not semantics that allowed participants to learn the grammar, but it was the presence of cues marking syntactic boundaries that facilitated learning in Moeser and Bregman’s experiment. In their experiment Morgan and Newport tested separately the role of the grouping cue marking the constituency, and the role of semantics. They showed that participants provided with the constituent grouping cue successfully learned the grammar (accuracy about 84%) and their performance did not significantly differ from that of participants provided with additional semantic cues (accuracy 90%). These results supported the hypothesis that when cues marking the constituent organization of the sentences are available in the input, participants successfully learn a complex grammar. However, since a difference was observed in the learning trends of the two groups of participants (constituent grouping cues condition and syntax-semantics correlation condition), suggesting different learning strategies, these results do not completely rule out the possibility of a facilitatory role of semantics.

We observed that the results collected by Morgan and Newport (1981) and the results from Experiment 5 invalidated Moeser and Bregman’s (1972) conclusion about the necessity of semantics to learn a grammar. However, as we mentioned above, it is still possible that semantics plays a facilitatory role in acquisition. Experiment 6 aimed to investigate this hypothesis. Moreover, we can outline some differences between this experiment and Morgan

---

33 Since the grammars used by Moeser and Bregman (1972) and Morgan and Newport (1981) can be represented as finite state grammars, we consider that the term Phrase Structure Grammar is used in this context to indicate a grammar as defined by rewrite rules and not to indicate only a subset of grammars in Chomsky’s hierarchy (i.e., context-free grammars).
and Newport’s one. First, these experiments were conducted in different modalities. The sentences in our experiment were presented auditorily, while they were presented visually in Morgan and Newport’s experiment. Second, the types of knowledge investigated were different. In our experiment, besides memorization, we also tested generalization to a novel lexicon and to new syntactic structures. Morgan and Newport presented all the possible types of sentences generated by the grammar during the training and also the lexicon was the same both in the training and in the test phases. Therefore, the novelty in the test sentences was determined by the fact that the linear sequence of the words was new, but not the syntactic structure. Third, we implemented semantics using videos showing geometrical shapes performing actions, while they used combinations of visual features (e.g., shape, colour, orientation). Finally, the phrase structure grammar investigated in Morgan and Newport’s study, generated sentence in which there were only adjacent dependencies (e.g., a CP must occur before a F word; a D word cannot appear after a word C34), whereas in our study the hierarchical agreement rule generated both adjacent and non-adjacent dependencies.

To sum up, we hypothesize that the presence of semantics would improve the level of command of the hierarchical agreement rule. Therefore, we expect that participants would be able to successfully learn the rule and would show a better performance compared to participants of Experiment 5.

34 Single capital letters indicate word class (i.e., class A, class C, class D, class E, class F), paired capital letters indicate phrases (i.e., CP, AP, BP).
4.2. Experiment 6

4.2.1. Method

Participants

Twenty participants took part in the experiment. They were all monolingual French speakers, with ages between ranging 20 and 37, with no reported hearing, reading or language impairment. They were paid or they received course credit for their participation.

Materials

Auditory Stimuli

The words used in this study were identical to those used in Experiment 5. All the phonological and prosodic cues used in the present experiment were identical to those used in Experiment 5.

Visual Stimuli

Each sentence was associated with a short video. Each noun was associated with a coloured geometric shape (i.e., a square, a circle, an oval, a star) while verbs were associated with actions (i.e., moving or turning). Singular nouns were represented by one geometric shape (e.g., a square, a circle, etc.) while plural nouns were represented by a group of three identical geometric shapes (e.g., three squares, three circles, etc.). Examples of video frames are shown in Figure 21.
Chapter 4. The Role of Semantic Cues

Figure 21. Examples of video frames for different sentence structures. Arrows indicate the movement of the colour-matched geometrical shape.

Videos were produced with Apple Keynote and then converted to .avi files with Windows Movie Maker.

Procedure

The procedure was the same as in Experiment 5 except for the task used in the lexicon learning sessions. In the present experiment, participants performed a picture matching task during the lexicon test phases. The lexicon learning sessions were then structured as follows:

Lexicon training phase: Participants were presented with 12 trials. In each trial the singular and the plural versions of a noun were presented. Two nouns were introduced in this phase (i.e., meku/i, Rabu/i). Each trial had the following structure: a fixation cross was shown in the centre of the screen for 500 ms, subsequently, the video was presented followed by a 500 ms blank screen interval which separated the trials. The videos in this phase were characterized as follows: the singular and the plural version of each noun were presented in one video which lasted 6.56 seconds. The singular version of the noun (or the plural one) was presented auditorily and after 560 ms the corresponding image appeared on the left side of the screen.
After 880 ms the other version of the noun (i.e., plural if the singular one was presented first, or singular if the plural version was presented first) was presented auditorily and the corresponding image appeared after an interval of 560 ms on the right side of the screen. The side of the screen where singular and plural nouns were presented was counterbalanced. Each noun was repeated six times in its singular form and six times in its plural form. Trials were presented in randomized order. Participants were asked to listen to the words and watch the images carefully.

**Lexicon test phase:** Participants were asked to perform a picture matching task. They were presented with 24 trials. Each trial had the following structure: a fixation cross was shown in the centre of the screen for 500 ms, subsequently, the video was presented followed by a 500 ms blank screen interval which separated the trials. The videos in this phase were characterized by the following features: they lasted 4.56 seconds and two images were on the screen at the beginning of the video. After 800 ms a noun was presented auditorily. One of the images corresponded to the word presented auditorily, while the other image could represent either the same noun with a different number or a different noun with the same number. Participants had to press the button of the response box that corresponded to the side of the screen where the image matching the noun presented auditorily was shown. If the image appeared on the left hand side they had to press the button labelled “G” (first letter of “Gauche”, which means “left” in French) whereas if the image appeared on the right hand side of the screen, they had to press the button labelled “D” (first letter of “Droite”, which means “right” in French).

**4.2.2. Results**

*D-prime Analysis*

Before performing the statistical analyses we eliminated the outliers. A response was defined as an outlier if the reaction time was two or more SD from the subject’s mean for each sentence’ length (2 words; 3 words; 4 words). As a result of this procedure, 5% of the responses were dropped.
Lexicon Learning

All participants performed well on the lexicon learning tests. The mean percentage of correct responses on the picture matching task was 95% ($SD = 5.2$) for the first lexicon test and 97% ($SD = 5.3$) for the new lexicon test. Performance did not differ significantly between the first and the second test ($z = 1.241, N – Ties = 15, p = .215$, two-tailed).

Performance with respect to chance level

Mean $d$-prime values for each type of knowledge (i.e., Memorization, Lexical Generalization, Lexical and Syntactic Generalization) and syntactic structure are summarized in Table 30.

Table 30. Mean $d$-prime values (SD in parenthesis) for each type of knowledge and sentence structure.

<table>
<thead>
<tr>
<th>Type of Knowledge</th>
<th>NV</th>
<th>N_NV</th>
<th>NN_V</th>
<th>NV_NV</th>
<th>NVN_V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memorization</td>
<td>5.02</td>
<td>0.93</td>
<td>1.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.34)</td>
<td>(2.04)</td>
<td>(2.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexical Generalization</td>
<td>2.13</td>
<td>1.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.37)</td>
<td>(2.53)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexical and Syntactic</td>
<td></td>
<td></td>
<td>2.11</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>Generalization</td>
<td></td>
<td></td>
<td>(2.47)</td>
<td>(1.47)</td>
<td></td>
</tr>
</tbody>
</table>

In order to examine whether participants had learned the rule we performed several one-sample t-tests$^{35}$ comparing the mean $d$-prime values to the chance level (i.e., as previously mentioned, zero is the chance level for $d$-prime values).

Memorization – Performance was significantly better than chance for the NV structure ($t(19) = 16.692, p < .001$), NN_V structure ($t(19) = 3.164, p = .002$), and N_NV structure ($t(19) = 2.046, p = .027$).

$^{35}$ The use of t-tests is justified by the fact that the distribution of $d$-prime values meets the assumption of normality ($p > .05$ for Kolmogorov-Smirnov tests) for each type of structure.
Lexical Generalization – Performance was significantly more accurate than chance for the NN_V structure ($t(19) = 3.177$, $p = .002$) and also for the N_NV structure ($t(19) = 4.020$, $p < .001$).

Lexical and Syntactic Generalization – Performance was significantly above chance for the NV_NV structure ($t(19) = 3.821$, $p < .001$) and for the NVN_V structure ($t(19) = 1.984$, $p = .031$).

Performance as a function of the manipulated variables

We performed an ANOVA on d-prime values in order to study the effects of the type of knowledge and of Locality. The main effect of Type of knowledge ($F(2,38) = 2.429$, $p = .102$) and the main effect of Locality ($F(1,19) = 2.802$, $p = .111$) were not significant. There was a significant interaction between the two variables, $F(2,38) = 3.987$, $p = .027$. This interaction reflected the fact that only in the Lexical and Syntactic Generalization, the performance of the participants was significantly better for the sentences with the adjacent dependency ($M = 2.1$, $SD = 2.46$) than for the sentences with the non-adjacent dependency ($M = 0.65$, $SD = 1.47$), $t(19) = 2.800$, $p = .011$.

Template Analysis

The templates considered for this experiment were the same as the ones considered for Experiment 3.

Three-Word Sentences

As shown in Figure 22, nine participants had their highest d-prime value for the hierarchical subject-verb agreement rule and they performed above chance. The performance of eleven participants was not above chance for any of the templates.
Figure 22. Individual participant performance for the three word sentences as a function of the template with the highest d-prime value. The line represents the threshold above which a performance was considered above chance.

Four-Word Sentences

As shown in Figure 23, nine participants had their highest d-prime value for the hierarchical subject-verb agreement rule and they performed above chance. Six of these participants performed above chance also for the three-word sentences, while two of them performed slightly below chance for the three-word sentences. Three participants performed above chance with the third template (i.e., $3^\wedge-4^\wedge$). The performance of eight participants was not above chance for any of the templates.
**Chapter 4.** The Role of Semantic Cues

**Figure 23.** Individual participant performance for four words sentences as a function of the template with the highest d-prime value. The line represents the threshold above which a performance was considered above chance.

**Reaction Times Analysis**

Mean RTs for each type of knowledge (i.e., Memorization, Lexical Generalization, Lexical and Syntactic Generalization) and syntactic structure are summarized in Table 31.

<table>
<thead>
<tr>
<th>Type of Knowledge</th>
<th>NV</th>
<th>N_NV</th>
<th>NN_V</th>
<th>NV_NV</th>
<th>NVN_V</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Memorization</td>
<td>495</td>
<td>1006</td>
<td>1046</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(221)</td>
<td>(461)</td>
<td>(511)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexical Generalization</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>971</td>
<td>1070</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(478)</td>
<td>(662)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexical and Syntactic Generalization</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1005</td>
<td>1121</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(501)</td>
<td>(602)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We performed an ANOVA on reaction times in order to study Type of knowledge and Locality effects. There was a significant main effect of Locality, $F(1,19) = 4.503$, $p = .047$. 

183
Participants were faster in judging adjacent dependencies ($M = 994$ ms, $SD = 472$) than non-adjacent dependencies ($M = 1079$ ms, $SD = 585$). The main effect of Type of knowledge and the interaction between the two variables were not significant ($F_s < 1$).

### 4.2.3. Discussion

The main purpose of the present experiment was to study the contribution of semantics to the learning of a hierarchical agreement rule. We consider first the pattern of results of the t-tests comparing participants’ performance to chance level. We observed that participants performed above chance for all types of knowledge (i.e., Memorization, Lexical Generalization, Lexical and Syntactic Generalization) and for all sentence structures investigated. This indicates that participants were able to learn the rule and generalize this knowledge both to a novel lexicon and to new syntactic structures.

The results of the analysis of variance showed in addition that the advantage for the adjacent dependencies over non-adjacent dependencies was present for Lexical and Syntactic Generalization but absent for Memorization and Lexical Generalization. The only difference found between adjacent and non-adjacent dependencies—the one for the Lexical and Syntactic Generalization—might be due to the fact that the structure NVN_V is more difficult compared to NV_NV. In the previous experiments we indeed observed that the NV_NV structure always had an advantage with respect to the other sentence structures, possibly because it is processed as two independent NV structures.

Also in this experiment we can identify four different profiles of performance, as in Experiment 5. 1) The first profile refers to participants who did not learn the rule or any other alternative template. 2) The second refers to participants who learned an alternative template. The analysis of the four-word sentences revealed that there are three participants who employed the template 3^4. 3) Participants belonging to the third profile learned the rule and generalized it only to the novel lexicon. 4) Finally, participants belonging to the fourth profile learned the rule and generalized it both to the novel lexicon and to new syntactic structure. Interestingly, two participants, performing slightly under the threshold of significance with the rule for the three-word sentences, improved their performance for the four-word sentences.
crossing the threshold and showing generalization to a novel lexicon and new syntactic structures. The question that arises from this observation and that will be addressed in the General Discussion concerns the nature of the knowledge that these participants acquired.

4.3. Comparisons between Experiment 5 and Experiment 6

4.3.1. Results

D-prime Analysis

Mean d-prime values for each type of knowledge, sentence structure and experiment are represented in Figure 24.

![Figure 24](image)

Figure 24. Mean d-prime values for each type of knowledge and sentence structure common to Experiment 5 and Experiment 6 (Error bars represent half standard deviation).

We performed a mixed ANOVA on the d-prime values; the between-subject variable was the Experiment (i.e., Experiment 5, Experiment 6); the two within-subject variables were Type of knowledge and Locality. The Greenhouse-Geisser Epsilon was used in order to deal with lack of sphericity for the factor Type of knowledge. The analysis revealed a main effect of the Locality: $F(1,38) = 12.640, p = .001$. Performance was better for sentences with an adjacent dependency ($M = 1.60, SD = 2.05$) than for sentences with a non-adjacent dependency ($M = 0.97, SD = 2.02$). The main effects of Experiment and Type of knowledge
were not significant ($F_s < 1$). The interaction between Type of knowledge and Locality was significant: $F(2,1.98) = 5.814, p = .005$. Pairwise comparisons revealed that performance was significantly better for sentences with an adjacent dependency than for sentences with a non-adjacent dependency only in the Lexical and Syntactic Generalization, ($t(39) = 4.613, p < .001$). No differences were found for Memorization ($t < 1, ns$) or for Lexical Generalization, ($t(39) = 1.229, p = .226$). The interaction between Locality and Experiment ($F(1,38) = 1.265, p = .268$), the interaction between Type of knowledge and Experiment ($F(1.692,76) = 1.997, p = .151$), and the interaction between Type of knowledge, Locality and Experiment ($F < 1$) were not significant.

**Reaction Times Analysis**

We performed a mixed ANOVA on the reaction times; the between subject variable was the Experiment (i.e., Experiment 5, Experiment 6); the two within subject variables were Type of knowledge and Locality. The Greenhouse-Geisser Epsilon was used in order to deal with lack of sphericity for the factor Type of knowledge. The analysis revealed a main effect of the Locality: $F(1,38) = 3.903, p = .055$. Sentences with an adjacent dependency were judged faster ($M = 974$ ms, $SD = 446$) than the ones with a non-adjacent dependency ($M = 1030$ ms, $SD = 513$). The main effects of Experiment and of Type of knowledge were not significant ($F_s < 1$). The interaction between Type of knowledge and Locality and the interaction between Type of knowledge, Locality and Experiment were not significant ($F_s < 1$). Also the interaction between Locality and Experiment was not significant ($F(1,38) = 1.096, p = .302$).

**4.3.2. Discussion**

The comparison between Experiment 5 and Experiment 6 was meant to clarify the contribution of semantics to the learning of a hierarchical agreement rule. The hypothesis we tested was whether semantics significantly improved artificial language learning. The results

---

36 Another mixed ANOVA has been performed in order to disentangle the role of semantics in memorization and in generalization. Two arguments motivated this choice: first, memorization and generalization are different cognitive processes; second, it is possible that semantics makes memorization difficult, since there is more information to track. Therefore, we performed a mixed ANOVA on d-prime values considering only two types of knowledge (i.e., Lexical Generalization, Lexical and Syntactic Generalization). However, also in this case there was not a significant main effect of the variable Experiment ($F(1,38) = 2.169, p = .149$).
from the analysis of variance on d-prime values pointed out that there is no significant
difference between participants’ performance to Experiment 5 and Experiment 6. This
observation suggests that semantics does not improve the learning of the hierarchical
agreement rule in our artificial language.

However, if we look at the results of the t-tests comparing participants’ scores to
chance level for each type of knowledge and sentence structure, one difference emerges. In
Experiment 5 only the performance for the NV_NV structure in the Lexical and Syntactic
Generalization was significantly above chance while in Experiment 6, the performance for
both sentence structures (i.e., NV_NV and NVN_V) was significantly above chance. The
present results indicate that learning is good in both experiments but it seems that the presence
of semantics increases the generalization of the rule to a novel lexicon and new syntactic
structure in the most difficult sentence structure. We can speculate that semantics played a role
only when a higher level processing was required compared to simple memorization.
Interestingly, when we considered only the Lexical Generalization and the Lexical and
Syntactic Generalization conditions in the analysis of variance, the $p$ value of the main effect
of the variable Experiment was not extremely high ($p = .149$). It is reasonable to assume that
the lack of statistically significant difference between the two experiments was due to the high
variability in the groups’ performances. Therefore, we can think that an increase of statistical
power (i.e., number of participants) would lead to a significant difference between
experiments with an advantage for the one involving semantics.

Since Lexical and Syntactic Generalization is tested with four-word sentences, we can
obtain additional information about the differences between Experiment 5 and Experiment 6
by looking at the template analysis. In Experiment 6, nine participants performed above
chance with the rule, while in Experiment 5 only five participants performed above chance
with the rule. This observation shows that there was an increment in the number of
participants successfully learning the rule for this type of knowledge. Moreover, in
Experiment 6 only three participants performed above chance with an alternative template,
while in Experiment 5 they were six. This observation suggests that fewer participants might
have misled and applied an alternative rule. Therefore, we can reasonably conclude that there
is an influence of semantics when the learning involves a higher level type of knowledge, such
as the one allowing the generalization of the rule both to a novel lexicon and new syntactic structures.

To sum up, the present results suggest that in order to learn the hierarchical agreement rule, the combination of cues marking agreement and constituency is sufficient and that semantics plays only a marginal role. However, the analysis of individual performances to the experiments suggests that there was an increase in the number of participants learning the rule when semantics was added. The results of the present experiment strongly raise the question of what is ultimately learned in an ALL paradigm, and which type of knowledge can account for a performance that is statistically above chance but it is far from being perfect. This issue will be developed in the next chapter.
Chapter 5. General Discussion
5.1. Summary of the Dissertation

The purpose of this thesis was twofold. First, we wanted to collect evidence about adults’ capacity to learn agreement rules in the context of an AL; second, we wanted to investigate the factors influencing this learning. We focused on the learning of two types of agreement dependencies: linear dependencies (i.e., based on the position of the elements in the linear sequence) and hierarchical dependencies (i.e., based on the position of the elements in the constituent structure).

While the ability to learn linear syntactic dependencies, such as Finite State Grammars, in the context of AL has been demonstrated by several studies (see for example Gomez & Gerken, 1999; Reber, 1967, 1989 among many others), conflicting evidence has been collected on the capacity to learn hierarchical syntactic dependencies (e.g., Phrase Structure Grammars) in artificial languages (e.g., Bahlmann et al., 2008; Fitch & Hauser, 2004; Mueller et al., 2010; Perruchet & Rey, 2005; de Vries et al., 2008). The interest of investigating these types of structures comes from the idea that, according to the Universal Grammar Theory (see Section 1.2.1.1.), one of the principles characterizing all natural languages is that linguistic information is processed and mapped in hierarchical structures rather than in linear sequences of words (Chomsky, 1965, 1995). Therefore, learners of an artificial language should spontaneously derive a hierarchical syntactic representation from the stimuli they are exposed to, rather than a linear representation. We decided to study the acquisition of an artificial rule mimicking the subject-verb agreement rules in number that we find in natural languages. In one case the rule was characterized by linear dependencies and in the other by hierarchical dependencies.

The investigation of the linear subject-verb agreement rule (see Chapter 2) allowed us to explore the learning of adjacent and non-adjacent dependencies based on the fixed position of the elements in the sequences, and the influence of the type of units (tone-words vs. pseudo-words) employed in the AL. The investigation of the hierarchical subject-verb agreement rule (see Chapter 3) allowed us to propose a novel type of rule that did not follow the commonly used distinction between Finite State Grammars and Phrase Structure Grammars. We
implemented an agreement rule based on the position of the subject and the verb in the constituents building up the sentences of the artificial language. We did that by marking grammatical categories with phonological cues and constituents’ boundaries with prosodic cues. Prosodic cues (i.e., pauses and pitch) were meant to indicate the boundaries of the constituents marking the hierarchical structure of the sentences. We explored the influence of different types of phonological cues, such as place of articulation, number of syllables, final vowels, as grammatical categories and morphological agreement markers on learning.

We also decided to investigate the influence of semantic cues on learning (see Chapter 4). A link between syntax and semantics has been proposed in accounts of syntactic processing and language acquisition. Modular models claim that semantic information does not play an immediate role in syntactic processing, while interactive models claim that semantic and context information play a role early in the syntactic processing of sentences (see Section 1.3.2). Semantics seems to play a role in language acquisition. In fact, the semantic bootstrapping hypothesis proposes that during language acquisition learners may rely on the correspondence between semantic and syntactic information to learn properties of their language such as syntactic categories. In the literature only a few studies directly explored the role of semantics in artificial language learning and reached contrasting conclusions. We decided to investigate whether by introducing semantic cues in our artificial language we could facilitate the processing and the learning of the hierarchical rule (Experiment 6).

All experiments included training and test phases. In the training phases participants were presented with sentences generated by the rule investigated, and in the test phase they were asked to perform a grammaticality judgment task in which they had to determine whether the sentences presented were generated by the rule or not. As different types of test materials were used, we were able to study different types of knowledge. Namely, memorization (when the grammatical sentences were the same as in the training phase), lexical generalization (when the grammatical sentences had a different lexicon compared to those in the training phase, but the same syntactic structures), syntactic generalization (when the grammatical sentences had different syntactic structures compared to those in the training phase, but the same lexicon), and lexical and syntactic generalization (when the grammatical sentences had different syntactic structures and a different lexicon compared to those in the training phase).
In this final chapter we will discuss the main results of the experiments further and their relevance in the debates about language acquisition and artificial language learning. We will highlight limitations of the empirical work and propose directions for future research.

5.2. Are adults able to learn linear dependencies in an artificial language?

In the first two experiments presented in this thesis we investigated whether adults were able to learn a linear subject-verb agreement rule in two different domains, music and language. This rule (i.e., the verb agrees with the first noun in the sentence) was based on the position of the elements (i.e., tone-words or pseudo-words) in the string (i.e., sentence). The results showed that participants did not learn the rule in either domain. In particular, in the music domain participants were not able to learn any regularity or alternative template. In the linguistic domain, participants could be classified into two groups according to their performance. Participants belonging to the first group did not learn the agreement rule or any alternative template, while participants belonging to the second group learned one of the alternative templates (i.e., the first, the second and the third elements have the same number). We assume that the perceptual salience of a particular category of items (i.e., grammatical matching sentences) misled those learners into tracking the wrong regularities.

We may wonder whether participants did not learn this rule since linear dependencies are not found in natural languages. However, we know from previous research that linear syntactic dependencies are learnable in artificial languages constituted by pseudo-words (see Gomez & Gerken, 2000 for a review) and also by other types of units, such as tones (e.g., Creel et al., 2004; Endress, 2010). We can identify two main differences between our experiments and previous research which may explain these conflicting results: the fact that our rule generated both adjacent and non-adjacent dependencies in the same language while previous rules did not, and the complexity of the units involved in the dependency which was greater in our study. We discuss them in turn.

While in previous research participants were asked to learn only one type of dependency (adjacent or non-adjacent) at a time, in Experiments 1 and 2 the linear agreement
rule created both adjacent and non-adjacent dependencies. Participants had therefore to be able to track both types of dependencies simultaneously. According to attention-based models (see Section 1.4.2.3.) the type of dependency learned in an ALL paradigm depends on the focus of participants’ attention during processing (Pacton & Perruchet, 2008). In ALL paradigms, attention is usually manipulated through instructions introducing the training phases. In our experiments participants were asked to perform a detection task - not directly linked to either the adjacent dependencies or the non-adjacent dependencies - (Experiment 1) or to passively listen to the list of sentences presented (Experiment 2). Therefore, we did not favour one type of dependency over the other. It is possible that the reason why participants did not abstract the rule is that they could not keep track of both types of linear dependencies at the same time.

Another difference between the experiments in the literature and the experiments in this dissertation (Experiments 1 and 2) is that the units involved in our experiments were more complex. Usually in the music domain, dependencies were determined over single tones, while in Experiment 1 they were determined over groups of three tones (tone-words). In the linguistic domain ALL studies employed monosyllabic or bisyllabic pseudo-words while the lexicon in Experiment 2 was built up of trisyllabic pseudo-words. Participants had to remember longer words and consequently more complex sequences. Memory load might have been higher in our experiments compared to previous studies using single tones or monosyllabic words. Interestingly, it has been recently suggested (M. C. Frank & Gibson, 2011) that memory constraints may play a role in artificial language acquisition. The authors proposed that memory limitations could explain previous negative findings in the learning and generalization of rule patterns (e.g., Endress et al., 2005; Gomez, 2002; Smith, 1966).

In Experiments 1 and 2 we employed different types of units building up the artificial languages (i.e., tone-words or pseudo-words), since we aimed to explore whether this factor had an influence on learning. However, because neither experiment revealed learning, it is hard to say whether the types of units played any role in the processing of the sentences.
5.3. Are adults able to learn hierarchical dependencies in an artificial language?

There has been debate about what grammar should be used to study hierarchical syntactic dependencies in ALL experiments. Most studies exploited centre-embedding grammars: A"B" (Fitch & Hauser, 2004; Friederici, 2004; Hauser, Chomsky, & Fitch, 2002) (see Section 1.4.2.1.2.). There is no consensus on the learnability of these grammars. It has been claimed that instead of acquiring the rule generating the centre-embedded sentences, learners apply simpler strategies, such as counting the A elements and comparing them to the number of B elements, to discriminate between grammatical and ungrammatical items (Hochmann, Azadpour, & Mehler, 2008; Perruchet & Rey, 2005; Pinker & Jackendoff, 2005; Zimmerer, 2010; de Vries et al., 2008). Therefore, this grammar seems not to be suitable to be used to investigate hierarchical rule learning, and other types of grammars generating hierarchical structures need to be investigated.

Subsequently, the application of Lindenmayer Grammars to the study of human linguistic abilities has been recently proposed by Saddy (2009) (see Section 1.4.2.1.2.). In particular, he exploited the learning of a Fibonacci Grammar: $A \rightarrow B \ B \rightarrow AB$ which generates strings like: A, B, AB, BAB, ABBAB, etc. In his study, after a short training phase, participants had to discriminate grammatical and ungrammatical sequences. The fact that participants were indeed able to discriminate sequences generated by the Fibonacci Grammar from other sequences suggested that they were able to extract the hierarchical structure that characterized the grammatical items. The abstraction of the structure-dependent regularities from the input was the only way to discriminate grammatical from ungrammatical sentences. It has been suggested that the computational complexity generated by Fibonacci grammar is “arbitrary large” (Piattelli-Palmarini & Uriagereka, 2008, p. 222) and if natural languages rely on this kind of system, then also linguistic computations would be “arbitrary large”. Constraints, such as attention or memory limitations, are needed to restrict this variability.

Our study develops within the framework of this debate. As a matter of fact, one of the aims of our work was to define a novel type of rule closer to those we find in natural language, which would allow the investigation of hierarchical syntactic dependencies learning (see
Section 1.5.1.). The hierarchical subject-verb agreement rule we created (i.e., the verb agrees with the highest noun in its constituent; if there is no noun in its constituent, the verb agrees with the highest noun in the immediately preceding constituent), was based on the representation of subject and verb positions in the constituents’ structure of the sentences. We employed phonological and prosodic cues inspired by those available in natural language to mark the morphological agreement and the constituents’ organization in the different types of structures generated by this rule.

The results obtained in Experiments 3 to 6 are highly heterogeneous; as shown in Table 32, participants were not equally successful in all learning steps and for all types of knowledge. In particular, we did not observe learning of the rule in Experiment 3, in which the lexicon was constituted by bisyllabic pseudo-words and in which categories and number were marked with different consonants and vowel phonetic features. In this experiment, participants were able to learn: the distinction between grammatical categories, and the association between agreement markers. In Experiment 4, in which the lexicon was constituted by monosyllabic pseudo-words and in which categories and number were marked with different consonant features, we observed learning of the rule but no generalization to new lexicon and novel syntactic structures. In Experiments 5 and 6, in which categories were cued by a different number of syllables and number by the final vowels, the group performance suggested that participants were able to learn and generalize the rule at a good mastery level. Looking more attentively at the individual performance, as we suggested in Section 1.5.3.2., we observed the existence of different participant profiles. At one extreme we had (also in these experiments) participants who did not manage to learn the rule and at the other extreme participants who were able to successfully learn the rule. Alternatively, some participants learned one of the alternative templates, and other participants performed well on one type of knowledge (e.g., Lexical Generalization) but not on another one (e.g., Lexical and Syntactic Generalization).
Table 32.
Main results for Experiments 3 to 6 for Learning Steps (see Chapter 3) and Generalization. “X” symbols indicate no learning, “√” symbols indicate learning, “~” symbols indicate that generalization was limited and not found for every type of knowledge investigated.

<table>
<thead>
<tr>
<th>Experiments</th>
<th>Grammar categories &amp; number</th>
<th>Association agreement markers</th>
<th>Hierarchical Agreement Rule</th>
<th>Generalization of the Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 3</td>
<td>√</td>
<td>√</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Experiment 4</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>~</td>
</tr>
<tr>
<td>Experiment 5</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Experiment 6</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>

The general conclusion we can draw from these results is that adults are able to learn hierarchical syntactic dependencies in AL. In fact, if there is at least a small subset of human adults able to learn this agreement rule then human adults are able to learn hierarchical syntactic dependencies. Since in Experiments 5 and 6 there were participants who acquired the rule and also generalized their knowledge of the rule to novel sentences, we can answer “yes” to our first research question. Therefore, the main question is no longer whether adults are able to learn hierarchical syntactic dependencies, but what are the factors that influence learning. This conclusion has important implications for the theories on language acquisition. In fact, while generativist theories claim that learners are biased towards hierarchical representations of the linguistic input (i.e., acquire and process sentences as hierarchically structured), usage-based theories state that learners are biased towards linear representations of the sentence structures. Therefore, the latter interpretation suggests a general preference in processing the linear dependencies of the sentences. At first glance, our data seem to suggest a bias towards hierarchical representation of the sentence structure, since some participants showed learning for the hierarchical rule but none for the linear one. In light of these results it would be tempting to conclude in favour of the generativist theories’ claims. However, the variables manipulated in these experiments were not the same, which makes the comparison between the learning of the rules impossible. Future work should attempt to perform a more strict comparison between the acquisitions of these two types of dependencies.
5.3.1. Phonological and Prosodic Cues

In Experiments 3, 4, and 5 we employed different types of phonological and prosodic cues. In Experiments 3 and 4, grammatical categories and number were marked by different places of articulation and by the voicing of the consonants contained in the (last) syllables of the words. In Experiment 5 (and consequently in Experiment 6), grammatical categories were marked by a different number of syllables, and number was marked by the final vowels.

One of the conclusions we can draw from the present results is that there are phonological cues more effective than others in marking information such as grammatical categories and number. For instance, we observed that the performance in Experiment 5, in which the grammatical categories were marked by the number of syllables, improved compared to Experiment 4, in which the grammatical categories were cued by the place of articulation of some phonemes in the words. We can speculate that the use of the number of syllables to categorize words is more efficient than place of articulation of the phonemes. This is not surprising, since the difference in syllable number is one of the phonological properties correlating with grammatical categories (Kelly, 1992, 1996). It has also been shown that newborns are already able to discriminate utterances on the basis of the number of syllables (Bijeljac-Babic, Bertoncini, & Mehler, 1993).

Most importantly, we observed that the performance in the learning task significantly improved when the morphological agreement was cued by the identity of the subject and verb last vowels, rather than when it was cued simply by the co-occurrence of syllables. This observation is relevant for theoretical models of natural language acquisition. In fact, it can be interpreted as evidence of the special status of the identity relationship. This status is recognized in particular by the perceptual primitives approach to language acquisition. According to this approach (see Section 1.4.2.3.), a perceptual primitive is a mechanism, originally non-linguistic, that is recruited by the language faculty and acts as filter determining which regularities in the input are learnable and which are not. One of the primitives that have been identified is characterized as a repetition-based mechanism. The presence of repetitions in the grammar triggers its learning. The improvement in rule learning when the morphological agreement was based on the repetition of the same final vowels for the subject
and the verb may have been determined by the fact that human adults can exploit this perceptual primitive which guides the acquisition of complex syntactic systems in our artificial language, and reasonably guides the acquisition and the processing of natural language as well. In fact, as Gervain pointed out (2008), both infant-directed and adult-directed speech contains repetitions which are supposed to facilitate natural language acquisition and processing. Repetitions are characterized by different levels of abstraction. An example of repetition at the phoneme level is taken from inflectional morphology of a Central Philippine language (i.e., Bikol) in which repetition is used to form diminutives (aloy “time span”, aloy aloy “short time”). At a more abstract level, adults are able to determine that Mary and The daughter of the neighbours are both determiner phrases which may play the same syntactic role (e.g., subject) in the sentence. Further evidence is needed to investigate the role of the perceptual primitives in the learning of more complex rules and to create a plausible sample book of all characterizations of these mechanisms.

In what concerns prosodic cues we reached two conclusions. First, the improvement in participants’ performance from Experiment 3 to Experiment 4 may have benefited from the presence of the 25ms pauses between words. These pauses facilitated the segmentation of the signal and reduced the cognitive burden which might have resulted from the fact that in Experiment 3 participants had to perform two tasks at the same time (segmentation and rule learning). Facilitating segmentation ensured that more cognitive resources were available and dedicated to rule extraction. Second, using longer pauses between constituents improved the implementation of the hierarchical rule. In fact, learning to parse the sentence into constituents is a necessary condition for the abstraction of the hierarchical subject-verb agreement rule, as it constitutes the last of the steps leading to learning (see Chapter 3, step 3). Our results are in line with interpretations according to which the extraction of some types of syntactic dependencies depends on the presence of bracketing cues in the input (Morgan & Newport, 1981; Morgan et al., 1987).

5.3.2. Semantic Cues

The role of semantics in natural language processing has been extensively investigated (see Section 1.3.2.), however only a few studies directly explored this topic in artificial
languages (Moeser & Bregman, 1972; Morgan & Newport, 1981). We contributed to the collection of evidence in the latter domain comparing Experiments 5 and 6. In Experiment 6 we studied the influence of semantic cues creating a correspondence between pseudo-words indicating nouns and geometrical shapes, and between pseudo-words indicating verbs and actions. The comparison of the results of this experiment with results of Experiment 5 (in which the only difference was the absence of semantic cues) suggested that: 1) Semantic cues are not necessary as proposed by Moeser and Bregman (1972), since learning was already observed in Experiment 5 and there is no significant difference between the groups’ performances of the two experiments; 2) It is possible that semantic cues nevertheless influence learning, since there were more participants learning the rule and less learning templates in Experiment 6 than in Experiment 5.

Our conclusions are in line with those proposed by Morgan and Newport (1981) according to whom the *conditio sine qua non* for hierarchical structures learning is the presence of cues marking the constituents’ structure of the sentences, while the semantic cues to the representation of these dependencies are not necessary to improve learning. The factors influencing/constraining the learning of hierarchical syntactic dependencies have to be searched in the integration of multiple overlapping cues of different nature (e.g., statistical, prosodic, perceptual) that trigger structure computations (see Section 1.4.2.3.).

5.3.3. Summary

To sum up, the main conclusions that can be draw from the data reported are the following:

1) Human adults are able to learn hierarchical syntactic structures in an artificial language under certain circumstances.

2) The circumstances that favour learning are related to the choice of adequate phonological cues, namely the most salient ones (e.g., the cue based on the number of syllables is more effective than the cue based on the place of articulation of the consonants in the words).
3) The choice of the characteristics of the elements involved in the syntactic dependency has to be made keeping in mind that there are specialized mechanisms that may constrain learning (e.g., the identity relation has to be preferred to simple co-occurrence, since there may be a perceptual primitive dedicated to repetitions).

4) Pauses marking words’ boundaries favour the learning of hierarchical syntactic dependencies, since they facilitate the segmentation of the signal and ensure that more cognitive resources are dedicated to rule extraction.

5) Semantic cues are not necessary to learn the hierarchical agreement rule.

5.4. Nature of the Acquired Knowledge

As already acknowledged in the discussions of the results (see Sections 3.4.3. and 4.1.3.), it is difficult to characterize the nature of the knowledge that results from the artificial language learning paradigm employed in this thesis. This difficulty is due to the modesty of the performance and its heterogeneity. Nevertheless, we can advance two hypotheses, in our opinion equally plausible, to account for the nature of the knowledge acquired during our experiments. According to the first hypothesis, participants would have acquired an abstract knowledge of the rule. The presence of different types of “variables”, such as cognitive limitations or inadequate learning conditions, might have negatively influenced their performance. According to the second hypothesis, participants would have acquired a knowledge 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. These hypotheses are detailed in the following sections.

Abstract Rule Knowledge

We can speculate that participants who successfully judged the grammaticality of the sentences and successfully performed the generalization to a novel lexicon and to new syntactic structures had learned the agreement rule and had an abstract representation of this rule. In fact, in order to be able to generalize their knowledge to novel words and structures, the participants must have had representations of the rule that are surface-independent.
Participants must also have represented the hierarchical structure of the constituents, since they were able to assign the syntactic roles to the right words based on this organization. This characterization of the knowledge reflects the definition of “abstract representation” as proposed by Redington and Chater (1996), namely the knowledge of the deep structure of the input (see Section 1.4.2.5).

However, it is important to highlight the fact that, even when the performance of participants was significantly above chance, it was never perfect (never 100% correct responses). How can we explain this fact? In order to discuss this fact we have to introduce the distinction between linguistic competence and linguistic performance. According to the generativist approach, competence is the system of rules that determines the knowledge of linguistic processes (i.e., it determines what is acceptable and what is not in the speaker-listener language) and performance is the actual use of the competence, that is how language is actually used. This distinction is for example salient in the context of the learnability/usability of centre-embedded structures. In fact, as soon as there are more than two levels of embedding in a sentence, the sentence becomes very difficult to interpret (e.g., the man the girl the boy said saw left). The gap between competence and performance has been usually explained with the presence of performance system limitations, such as working memory limitations (Caplan & Waters, 1999; Chomsky, 1965; E. Gibson & Thomas, 1999). Long-distance dependencies, originated by embedding, are more difficult to interpret, since listeners have to store the information longer than for instance multiple tail-embeddings (e.g., the dog that bit the cat that ate the mouse).

It is possible that some of our participants had full competence of the agreement rule but the presence of other “variables” may explain why the performance was not successful 100% of the times. In other words, they may have developed full knowledge of the rule, but there were cognitive limitations influencing their performance. These limitations might be attention, the quantity and quality of information participants had to keep track of (i.e., grammatical categories, morphological agreement markers, relationship between the marker, sentence parsing, and syntactic categories), memory limitations, etc.
The modest performance might have also been determined by the fact that participants performed the learning task in inadequately learning conditions since they did not have the time to consolidate what they had learned. Training and test phases took place one right after the other and the experiments in this thesis usually lasted 60 minutes at most. It has been shown, for example, that memory performance and generalization are enhanced by sleep (e.g., Fenn, Nusbaum, & Margoliash, 2003; Walker & Stickgold, 2006). Therefore, we can assume that giving participants more time to perform the training and test in separate days may result in a better performance. Interestingly, it is possible that generalization would be successful only when the rule is better consolidated. For instance, in the context of categorization and perceptual skills it has been shown that generalization has a longer time course than learning and that they are subserved by partially distinct neural changes (Benard, Stach, & Giurfa, 2006; Wright, Wilson, & Sabin, 2010). This would explain also the profile of participants who were able to learn the rule but not to generalize their knowledge to a novel lexicon and/or new syntactic structures.

**Micro-Rule Knowledge**

Another hypothesis concerning the nature of the knowledge acquired during our study refers to the idea of micro-rule. We can speculate that the participants had acquired only some aspects of the target agreement rule and not its totality. In the literature, micro-rules are defined as rules that account only partially for the sentences generated by the target grammar and are the result of insights about fragments of the training materials (see Section 1.4.2.6.) (Dulany et al., 1984; Dulany, Carlson, & Dewey, 1985). For instance, instead of learning the finite-state grammar generating sequences like: MTSXS, MTPKK, etc., participants learn the micro-rule “M is always in first position”. The learning of these micro-rules does not lead to a good performance but can account for a performance better than chance that is typically obtained in AGL paradigms. Micro-rules are therefore dependent on the surface properties of the items.

In our experiments, we identified different templates that indeed can be considered as micro-rules. The fact that some of the participants applied one of the templates we proposed represents striking evidence that participants learned micro-rules. For instance, the template
which resulted the more likely to be learned for the four-word sentences in Experiment 3, 5 and 6 (i.e., the element in third position agrees with the element in fourth position) is a rule based on the knowledge of the morphological agreement markers but limited to the linear order of the words. Therefore, it could only partially account for the grammatical sentences generated by the hierarchical agreement rule.

In sum, it is possible that the modest success in our participants’ performance could be due to the fact that they actually did not learn the agreement rule but alternative micro-rules which did not imply any hierarchical representation of the sentences. This micro-rule could account for only a subset of grammatical sentences and, therefore, its knowledge would have never allowed participants to reach a perfect score at the learning task.

5.5. Methodological Issues

One of the aims of this thesis was to create an artificial language that was at the same time closer to natural language than the AL used in the literature, but also simple enough to be learnable in training sessions of a reasonable length. We decided to use an agreement rule since we wanted to draw a parallel between the learning of a similar rule in natural and artificial languages. The general idea was to compare error rates in the artificial language learning paradigm with attraction errors in natural language paradigms. We wanted to investigate whether an intervening element in an artificial language could attract agreement differently, depending on the nature and position of this element in the hierarchical structure of the sentence. Therefore, we chose different types of sentences characterized by adjacent and non-adjacent dependencies.

However, since participants were not able to successfully learn the linear agreement rule in Experiments 1 and 2, we decided to investigate the learnability of a hierarchical subject-verb agreement rule and the factors influencing participants’ performance. As we soon realized, there were a series of methodological choices that needed to be made which might have influenced the learning of the agreement rules. In the present section we will outline the main methodological issues that were raised in our investigation. These have to do with the following problems: the choice of the cues to mark hierarchical structure, the choice of the
syntactic structures, the choice of the training procedure, the choice of the test procedure, and the choice of the type of analysis.

**Choice of the cues to mark hierarchical structure**

A major difficulty we encountered in implementing the hierarchical agreement rule was how to cue the hierarchical structure of the sentences. We considered two possibilities: the use of clitics or prosody. In the first possibility the sentence structure could be implemented introducing a CV syllable (always the same) before the verb. In this case the agreement rule is characterized as follows: when the subject is plural then the verb is plural; however if the verb is prefixed, it is singular. We can represent this latter sentence as NxV were N is the subject, x is the clitic and V is the verb. The prefix would mimic the presence of a clitic intervening between the subject and the verb. However, we judged this type of cue problematic, since it could have created ambiguity with the sentence structure NNV. The “x” element could have been categorized as another noun (N). Therefore, we selected the second possibility, namely to use prosodic cues to mark constituents’ boundaries.

**Choice of the syntactic structures**

Once we determined the agreement rule characterizing our artificial language, there were multiple types of syntactic structures (i.e., sentences) generated by that rule. It was, therefore, necessary to decide which syntactic structures had to be included in the study. The sentence structures we chose were from two to four words long. This characteristic of the material allowed us to test participants’ learning starting from the easiest syntactic structures to the more complex ones. However, as we observed later on, the NV_NV sentence structure was easier to judge than we expected. In fact, it could have been analysed as two NV structures concatenated. Therefore, the syntactic generalization we observed for this structure was different from the one observed for the NVN_V structures, creating sometimes difficulties in the comprehension of the results. Nevertheless, we decided to keep the NV_NV sentences, since their presence allowed us to compare sentences with the same linear word order that were grammatical for one type of syntactic structure but not for the other (e.g., *lepu stu* *govi sti* vs. *lepu stu govi sti*). This comparison gave us the possibility to directly control that participants were not judging the sentences only based on differences in the linear order of the
words. In order to correctly discriminate the grammatical \([lepu \ stu] \ [govi \ sti]\) and the ungrammatical \(*[[lepu \ stu \ govi] \ sti]$$\), participants had to be able to represent the hierarchical structure of the sentences.

**Choice of the training procedure**

Another methodological issue concerns the choice of appropriate training. Most studies in ALL literature are characterized by passive training in which participants are asked to simply listen to lists of sentences generated by the target grammar. This choice has two possible limitations. First, the fact that learners are confronted with the passive listening (or visualization) of grammatical strings is quite different from the nature of the “training” the speaker-listeners receive during natural language acquisition. In fact, when infants learn their first language or adults learn a second language, they are immersed in a meaningful context (e.g., visual context, social context, etc.) which is not available in ALL studies. Second, training phases with passive listening may challenge participants’ capacity to maintain the attention through the whole experiment.

However, introducing an additional task during training to keep participants’ attention may have even worse consequences on learning. In fact, as we claimed in the discussion of Experiment 1 (see Section 2.1.3.), the use of a second task (in this case a detection task) may have prevented learning by directing participants’ attention to the wrong properties of the input. As presented in Section 1.4.2.3., attention is an important factor influencing learning. Directing participants’ attention towards some aspects of the materials may favour the learning of certain regularities and prevent the learning of others. Moreover, as it has been extensively studied in the context of dual-task interference, executing two concurrent tasks may negatively influence people’s performance (see Pashler, 1994 for a review).

**Choice of the test procedure**

There are some methodological issues related to the test phase that may be improved in future research. In particular, in order to facilitate the interpretation of the results of a complex learning task, it would be useful to test the learning process step by step. This suggestion will be easier to understand if we refer to the three learning steps characterizing the hierarchical
agreement rule learning (see Chapter 3). We assume that in order to successfully learn the rule, participants have to go through three steps: learn the categorization of the words, learn the morphological agreement markers, and learn the rule. In the present work we did not test learners’ performance at each step, but only the third step. This resulted in a loss of valuable information for the discussion of the experiments’ results. More specifically, we can only speculate about the participants’ success in each learning step.

However, it is important to underline that introducing many test phases could also have been problematic. In fact, it is possible that participants keep learning during the test as well and not only during training, making the learning process difficult to control. Testing participants on parts of the rule or precise aspects of the materials may turn their attention away from the general target grammar. For example, we observed in Experiment 4 that testing participants only on the NV structure with a novel lexicon might have prevented participants to generalize the rule to non-adjacent dependencies. Therefore, we suggest that future researchers pay attention to finding a good balance between investigating each step of the process leading to learning or only the final product.

**Choice of the type of analysis**

A fifth and final important methodological issue concerns the template analysis. This type of analysis had the great advantage of allowing us to better investigate and understand the individual differences between participants’ performance. By considering these differences and the alternative rules participants might have learned during the task, we can deepen our understanding of the type of knowledge resulting from artificial language learning. However, we encountered two difficulties in implementing this analysis. The first difficulty concerns the definition of the templates. There are several alternative rules that can fit more or less the materials, but we cannot consider all of them. How to choose? Which are the most plausible alternatives to the target rule? For instance, in Experiments 3 to 6 we chose the templates based on the relationship between morphological agreement markers, since this allowed us to compare the processing of linear and hierarchical dependencies. However, other possible templates could have been chosen.
The second difficulty concerns how to statistically calculate which template fits best the participants’ performance. In this thesis we fixed a threshold (d-prime = 1.00) above which the performance of a participant was considered as better than chance for that template, which was the one best characterizing the performance for a certain participant. To our knowledge there is only one other study that investigated individual behaviour in an ALL paradigm with a similar analysis (Zimmerer et al., 2010). However, due to methodological differences between this study and our work (e.g., in the test phase they had a different number of grammatical and ungrammatical items) we could not use the same principle to determine the significance threshold (which was based on the binomial test). Therefore, we opted for a fixed threshold, since it seemed a plausible way to proceed. Indeed this type of analysis needs to be further developed and improved, but it represents a valuable tool for future research in ALL.

5.6. Future Perspectives

This thesis only represents a first attempt to study an artificial language that is more “natural” than the ones usually investigated in ALL paradigms. Clearly more data are necessary to further detail the learnability of hierarchical syntactic dependencies and to characterize the representation and processing of these structures. In the present section we will outline some interesting perspectives for future research using artificial languages.

Further studies should explore the processing of hierarchical syntactic dependencies by infants and newborns. The interest of this investigation resides in the fact that evidence of learning very early in life could contribute to the debate about the innateness of learning mechanisms that might facilitate language acquisition. Several studies showed that infants are indeed able to learn adjacent and non-adjacent linear dependencies (e.g., Saffran, Aslin, et al., 1996; Saffran et al., 1999), and algebraic rules (Marcus et al., 1999). However, there are no studies that directly tested infants’ capacity to process hierarchical syntactic dependencies. Further research is needed to evaluate newborns and infants’ ability to acquire this type of dependencies in ALL experiments and to explore the nature of the learning mechanisms involved in this task.
In addition, the use of artificial languages could contribute to the understanding of learning processes used in second language acquisition. The idea of drawing a parallel between these two lines of research has gained interest in recent years (Carroll et al., 2009; Mueller, 2006; Saddy, 2009). In particular two possible research questions could be investigated: 1) Are the same neural networks activated during sentence processing in L2 and in artificial languages? 2) Is there transfer of L1 knowledge to L2 and artificial language learning?

Finally, artificial languages can be employed in neuroimaging studies exploring how complex rule learning could generate changes in brain activity. For instance, three different investigations could be carried out: 1) The comparison of the brain activation before and after the training with complex syntactic rules, such as the hierarchical subject-verb agreement rule proposed in this thesis. 2) The comparison of the brain activation of two groups of participants learning different types of rules, such as linear vs. hierarchical (e.g., Tettamanti et al., 2002, 2009) or two rules in which only one variable was manipulated (different prosodic cues, different phonological cues, or different modality). 3) Investigate how much exposure to an artificial language is required to observe a change in the brain activity. These investigations should provide insights into the open questions about rule learning.
REFERENCES


References


Psychology, 57(1), 56-74.


Artificial Grammars From Exemplars? *Cognitive science, 32*(6), 1021-1036.


Knowlton, B. J., & Squire, L. R. (1996). Artificial grammar learning depends on implicit


Cambridge, MA: The MIT Press.


Psychological Science, 8(2), 103-105.


List of Appendixes

APPENDIX A. Stimuli Experiment 1

APPENDIX B. Stimuli Experiment 2 (IPA Alphabet)

APPENDIX C. Stimuli Experiment 3 (SAMPA Alphabet)

APPENDIX D. Stimuli Experiment 4 (SAMPA Alphabet)

APPENDIX E. Stimuli Experiments 5 and 6 (SAMPA Alphabet)
# APPENDIX A. Stimuli Experiment 1

<table>
<thead>
<tr>
<th>NV</th>
<th>NVN</th>
<th>NVN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADB CC#D</td>
<td>ADB CC#D DFE</td>
<td>ADB DFE CC#D</td>
</tr>
<tr>
<td>ADB D#ED</td>
<td>ADB CC#D DFEg</td>
<td>ADB DFEg CC#D</td>
</tr>
<tr>
<td>ADB FCF#</td>
<td>ADB CC#D GG#A</td>
<td>ADB DFE D#ED</td>
</tr>
<tr>
<td>ADBg CC#Dg</td>
<td>ADB CC#D GG#Ag</td>
<td>ADB D#ED GG#A</td>
</tr>
<tr>
<td>ADBg D#EDg</td>
<td>ADB D#ED DFEg</td>
<td>ADB GG#A CC#D</td>
</tr>
<tr>
<td>ADBg FCF#g</td>
<td>ADB D#ED DFEg</td>
<td>ADB GG#Ag CC#D</td>
</tr>
<tr>
<td>DFE CC#D</td>
<td>ADB FCF# DFE</td>
<td>ADB GG#A FCF#</td>
</tr>
<tr>
<td>DFE D#ED</td>
<td>ADB FCF# GG#Ag</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFE FCF#</td>
<td>ADBg CC#Dg GG#A</td>
<td>ADB D#ED DFE</td>
</tr>
<tr>
<td>DFEg CC#Dg</td>
<td>ADBg CC#Dg GG#Ag</td>
<td>ADB D#ED DFE</td>
</tr>
<tr>
<td>DFEg D#EDg</td>
<td>ADBg D#EDg DFEg</td>
<td>ADB D#ED DFE</td>
</tr>
<tr>
<td>DFEg FCF#g</td>
<td>ADBg D#EDg DFEg</td>
<td>ADB D#ED DFE</td>
</tr>
<tr>
<td>GG#A CC#D</td>
<td>ADBg FCF#g GG#Ag</td>
<td>ADB D#ED DFE</td>
</tr>
<tr>
<td>GG#Ag D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg CC#Dg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#A CC#D</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg CC#Dg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#A CC#D</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg CC#Dg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#A CC#D</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg CC#Dg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#A CC#D</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg CC#Dg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#A CC#D</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg CC#Dg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#A CC#D</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg CC#Dg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#A CC#D</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg CC#Dg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#A CC#D</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg CC#Dg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#A CC#D</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg CC#Dg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#A CC#D</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg CC#Dg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#A CC#D</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg CC#Dg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#A CC#D</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg CC#Dg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#A CC#D</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag D#EDg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>GG#Ag FCF#g</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>DFEg CC#Dg</td>
<td>DFEg CC#Dg ADBg</td>
<td>ADB D#ED DFEg</td>
</tr>
<tr>
<td>Grammatical</td>
<td>Ungrammatical</td>
<td>Grammatical</td>
</tr>
<tr>
<td>-------------</td>
<td>---------------</td>
<td>-------------</td>
</tr>
<tr>
<td>ADB CC#D</td>
<td>*ADB CC#D</td>
<td>ADB D#ED</td>
</tr>
<tr>
<td>ADB D#ED</td>
<td>*ADB CC#D</td>
<td>ADB GD#ED</td>
</tr>
<tr>
<td>DFE CC#D</td>
<td>*DFE CC#D</td>
<td>ADB FCF#</td>
</tr>
<tr>
<td>DFE FCF#</td>
<td>*DFE FCF#</td>
<td>ADB CC#D</td>
</tr>
<tr>
<td>GG#A D#ED</td>
<td>*GG#A D#ED</td>
<td>ADB FCF#</td>
</tr>
<tr>
<td>GG#A FCF#</td>
<td>*GG#A FCF#</td>
<td>ADB FCF#</td>
</tr>
<tr>
<td>ADB CC#G</td>
<td>*ADB CC#G</td>
<td>ADB CC#D</td>
</tr>
<tr>
<td>ADB ED#G</td>
<td>*ADB ED#G</td>
<td>ADB CC#D</td>
</tr>
<tr>
<td>DFE CC#G</td>
<td>*DFE CC#G</td>
<td>ADB CC#D</td>
</tr>
<tr>
<td>DFE GCF#G</td>
<td>*DFE GCF#G</td>
<td>ADB CC#D</td>
</tr>
<tr>
<td>GG#A G D#ED</td>
<td>*GG#A G D#ED</td>
<td>ADB CC#D</td>
</tr>
<tr>
<td>GG#A G FCF#G</td>
<td>*GG#A G FCF#G</td>
<td>ADB CC#D</td>
</tr>
<tr>
<td>GG#A G FCF#G</td>
<td>*GG#A G FCF#G</td>
<td>ADB CC#D</td>
</tr>
</tbody>
</table>
## APPENDIX B. Stimuli Experiment 2 (IPA Alphabet)

<table>
<thead>
<tr>
<th></th>
<th>TRAINING</th>
<th>NNV</th>
</tr>
</thead>
<tbody>
<tr>
<td>NV</td>
<td>patobu vojusu</td>
<td>patobu vojusu tebipu</td>
</tr>
<tr>
<td></td>
<td>patobu syfepu</td>
<td>patobu vojusu tebipu</td>
</tr>
<tr>
<td></td>
<td>patobu fojepu</td>
<td>patobu vojusu bepstu</td>
</tr>
<tr>
<td></td>
<td>patobep vojus</td>
<td>patobu vojusu bepstu</td>
</tr>
<tr>
<td></td>
<td>patobep syfepu</td>
<td>patobu syfepu bepstu</td>
</tr>
<tr>
<td></td>
<td>patobep fojepu</td>
<td>patobu syfepu tebipu</td>
</tr>
<tr>
<td></td>
<td>tebipu vojusu</td>
<td>patobu fojepu bepstu</td>
</tr>
<tr>
<td></td>
<td>tebipu syfepu</td>
<td>patobu fojepu bepstu</td>
</tr>
<tr>
<td></td>
<td>tebipex vojus</td>
<td>patobep vojus bepstu</td>
</tr>
<tr>
<td></td>
<td>tebipex syfepu</td>
<td>patobep syfepu tebipu</td>
</tr>
<tr>
<td></td>
<td>bepstu vojus</td>
<td>patobep syfepu bepstu</td>
</tr>
<tr>
<td></td>
<td>bepstu syfepu</td>
<td>patobep syfepu bepstu</td>
</tr>
<tr>
<td></td>
<td>bepstu fojepu</td>
<td>patobep fojepu tebipu</td>
</tr>
<tr>
<td></td>
<td>bepsep vojus</td>
<td>tebipu vojusu patobu</td>
</tr>
<tr>
<td></td>
<td>bepsep syfepu</td>
<td>tebipu vojusu patobu</td>
</tr>
<tr>
<td></td>
<td>bepsep fojepu</td>
<td>tebipu syfepu patobu</td>
</tr>
<tr>
<td></td>
<td>tebipu syfepu</td>
<td>tebipu syfepu patobu</td>
</tr>
<tr>
<td></td>
<td>tebipu fojepu</td>
<td>tebipu fojepu patobu</td>
</tr>
<tr>
<td></td>
<td>tebipex vojus</td>
<td>tebipex vojus patobu</td>
</tr>
<tr>
<td></td>
<td>tebipex syfepu</td>
<td>tebipex syfepu patobu</td>
</tr>
<tr>
<td></td>
<td>tebipex fojepu</td>
<td>tebipex fojepu patobu</td>
</tr>
<tr>
<td></td>
<td>bepstu vojusu</td>
<td>bepstu vojusu patobu</td>
</tr>
<tr>
<td></td>
<td>bepstu syfepu</td>
<td>bepstu syfepu patobu</td>
</tr>
<tr>
<td></td>
<td>bepstu fojepu</td>
<td>bepstu fojepu patobu</td>
</tr>
<tr>
<td></td>
<td>bepsep vojus</td>
<td>bepsep vojus patobu</td>
</tr>
<tr>
<td></td>
<td>bepsep syfepu</td>
<td>bepsep syfepu patobu</td>
</tr>
<tr>
<td></td>
<td>bepsep fojepu</td>
<td>bepsep fojepu patobu</td>
</tr>
<tr>
<td></td>
<td>bepsep syfepu</td>
<td>bepsep fojepu patobu</td>
</tr>
<tr>
<td></td>
<td>bepsep fojepu</td>
<td>bepsep fojepu patobu</td>
</tr>
<tr>
<td></td>
<td>bepsep syfepu</td>
<td>bepsep fojepu patobu</td>
</tr>
<tr>
<td></td>
<td>bepsep fojepu</td>
<td>bepsep fojepu patobu</td>
</tr>
<tr>
<td></td>
<td>bepsep syfepu</td>
<td>bepsep fojepu patobu</td>
</tr>
<tr>
<td>Grammatical</td>
<td>Ungrammatical</td>
<td>Grammatical</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>patobu voľasu</td>
<td>'patobu voľasē</td>
<td>patobu syvenu těbipu</td>
</tr>
<tr>
<td>patobu syvenu</td>
<td>'patobu syveři</td>
<td>patobu syvenu bepōtě</td>
</tr>
<tr>
<td>těbipu voľasu</td>
<td>'těbipu voľasē</td>
<td>patobu fosev bepōtu</td>
</tr>
<tr>
<td>těbipu fosev</td>
<td>'těbipu fosevě</td>
<td>patobu fosev těbipě</td>
</tr>
<tr>
<td>bepōtu syvenu</td>
<td>'bepōtu syveři</td>
<td>těbipu voľasē těbipu</td>
</tr>
<tr>
<td>bepōtu fosev</td>
<td>'bepōtu fosevě</td>
<td>těbipu voľasē bepōtě</td>
</tr>
<tr>
<td>patobě voľasē</td>
<td>'patobě syveři</td>
<td>těbipu fosev patobu</td>
</tr>
<tr>
<td>patobě syveři</td>
<td>'patobě fosev</td>
<td>těbipu fosev patobě</td>
</tr>
<tr>
<td>těbipě voľasě</td>
<td>'těbipě voľasē</td>
<td>bepōtu fosev těbipu</td>
</tr>
<tr>
<td>těbipě fosev</td>
<td>'těbipě syveři</td>
<td>bepōtu fosev těbipě</td>
</tr>
<tr>
<td>bepōtě syveři</td>
<td>'bepōtě fosev</td>
<td>bepōtu fosev patobě</td>
</tr>
<tr>
<td>bepōtě fosev</td>
<td>'bepōtě fosevě</td>
<td>bepōtu fosev patobě</td>
</tr>
<tr>
<td>těbipě voľasě</td>
<td>'těbipě voľasē</td>
<td>bepōtu fosev těbipě</td>
</tr>
<tr>
<td>těbipě syveři</td>
<td>'těbipě syveři</td>
<td>bepōtu fosev těbipě</td>
</tr>
<tr>
<td>bepōtě syveři</td>
<td>'bepōtě fosev</td>
<td>bepōtu fosev těbipě</td>
</tr>
<tr>
<td>bepōtě fosev</td>
<td>'bepōtě fosevě</td>
<td>bepōtu fosev těbipě</td>
</tr>
</tbody>
</table>
### APPENDIX C. Stimuli Experiment 3 (SAMPA Alphabet)

<table>
<thead>
<tr>
<th>NV</th>
<th>N_NV</th>
<th>NN_V</th>
</tr>
</thead>
<tbody>
<tr>
<td>bladu fIEbo klyti vIEpa</td>
<td>bladu rItyi slepa</td>
<td>klydy tRityi vRapa</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fIEbo klyti vIEpa</td>
<td>bladu rItyi slepa</td>
<td>klydy tRityi vRapa</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
<td>bladu fIEbo gIEbo klydy dROdu vIEbo</td>
<td>bladu fOdu gIEbo klydy dROdu vIEbo</td>
</tr>
<tr>
<td>Grammatical</td>
<td>NV_NV</td>
<td>Ungrammatical</td>
</tr>
<tr>
<td>-------------</td>
<td>-------</td>
<td>---------------</td>
</tr>
<tr>
<td>kRodu bRobo fRyti slepa</td>
<td>*dROdu slepa_steti tRybo</td>
<td>blati plipavREdu _stypa</td>
</tr>
<tr>
<td>fRydu glRebo _vREti dRypa</td>
<td>*bRodu plypa_fOt iEbo</td>
<td>blati pRupavliti _kRypa</td>
</tr>
<tr>
<td>pROti plypa_pRidu stybo</td>
<td>*klydu bRupa_Ryti pRubo</td>
<td>bRoti slepavlydu _vRopa</td>
</tr>
<tr>
<td>bRoti kRypa_klidu vRobo</td>
<td>*dRudu stypa_fOt iEbo</td>
<td>dRidu vRaboklyti _klobo</td>
</tr>
<tr>
<td>sloti klopa_glEd u bRubo</td>
<td>*dRuti tRybo_sloti gRebo</td>
<td>dRutu slupasloti _kRypa</td>
</tr>
<tr>
<td>pRiti glapa_gRudu slebo</td>
<td>*kRodu glapa_pROdu pRupa</td>
<td>fLOdu slebosteti _glabo</td>
</tr>
<tr>
<td>kRoti dRypa_pRiti tRypa</td>
<td>*klyti stybo_vREti glabo</td>
<td>fLOti blopaklyti _flabo</td>
</tr>
<tr>
<td>vli du bblobo_steti fIEpa</td>
<td>*fRiti tRebo_gRudu vRapa</td>
<td>fRedi plybokRiti _fRObo</td>
</tr>
<tr>
<td>vlyti fIa pa_fIo ti plipa</td>
<td>*fRiti fROpa_dROdu klopa</td>
<td>fRedi vRopapRidi_dRypa</td>
</tr>
<tr>
<td>fReti fROpa_dROdu kRRebo</td>
<td>*kRuti fRObo_plibo slebo</td>
<td>glEti vIEpafi gRebo</td>
</tr>
<tr>
<td>stedu kRybo_oRodi glapa</td>
<td>*kRuti pRubo_blati tRybo</td>
<td>glEti vIEpafi gRebo</td>
</tr>
<tr>
<td>dRudu plybo klydu vIEbo</td>
<td>*stuti vIEbo_sledu tRapa</td>
<td>kRuti gRuidu _plibo</td>
</tr>
<tr>
<td>gIEd u kRRebo_pRidu pRubo</td>
<td>*fRedu kRrepagRadu plypa</td>
<td>kRudu blobofRydu _plibo</td>
</tr>
<tr>
<td>vli tu slupa_bEI Ei kRpa</td>
<td>*pledu flapa_tRidu vRapa</td>
<td>kRudu tRepapRati_dRypa</td>
</tr>
<tr>
<td>vREdu vRRebo_bi Ed u glabo</td>
<td>*fRyti plibo_pROti slubo</td>
<td>kRudu blobofRydu _plibo</td>
</tr>
<tr>
<td>klyti stypa_kli ti blopa</td>
<td>*gRati tRebo_gRebo flabo</td>
<td>kRudu blobofRydu _plibo</td>
</tr>
<tr>
<td>vlydu gIOb e_ple du dRybo</td>
<td>*sledu tRapa_sted u gIOpa</td>
<td>glEti vIEpa _pledu glObo</td>
</tr>
<tr>
<td>slodu fRObo_dRIdu vRabo</td>
<td>*vlydu fIEpa_fRe ti fRObo</td>
<td>pRidu stybosteti _gLObo</td>
</tr>
<tr>
<td>gRati slupa_tRiti gRepa</td>
<td>*studu vRopa _tRiti vRabo</td>
<td>pRidu bRupabRoti_fIEpa</td>
</tr>
<tr>
<td>tRibu klobo_dRuti gRepa</td>
<td>*bEti fROpa_bRiti flabo</td>
<td>pR0du dRybo gEd u_flabo</td>
</tr>
<tr>
<td>fIOt i bRupa_fRydu bRobo</td>
<td>*pROdu slepa_klyti blobo</td>
<td>siodu bRrobhobi tRipti _pRybo</td>
</tr>
<tr>
<td>gRuti kRypa_vREti bRopa</td>
<td>*steti vRabo_fOt iEbo</td>
<td>sloti dRypatRidu_tRypa</td>
</tr>
<tr>
<td>stuti bIopa_vlIdu klobo</td>
<td>*vIiti fIEbo_gRadox tRypa</td>
<td>stedu bRubopledu_flabo</td>
</tr>
<tr>
<td>kIdu plibo_fRed u bRubo</td>
<td>*bRoti gIOb e_gIEd u vIEbo</td>
<td>studu kRebo vlydu_fIEbo</td>
</tr>
<tr>
<td>gRudu sluba_bRobo fIEd u</td>
<td>*bIEd u bIoba_vREdu glapa</td>
<td>tRidu glabo dRudu _klobo</td>
</tr>
<tr>
<td>studu fIabo_gIEd u plipa</td>
<td>*vlyti pIoba_gRuti vIEbo</td>
<td>vlIdu klobolEt i _pRupa</td>
</tr>
<tr>
<td>blati kRapa_pleti tRapa</td>
<td>*fRiti kRybo_bRodu bRopa</td>
<td>vIiti bRopadRODu_pRupa</td>
</tr>
<tr>
<td>gRadu slebo_bEI Ei vRopa</td>
<td>*dROti pRubo_klydu pRupa</td>
<td>vREdu fIEbo gRudo_tRebo</td>
</tr>
</tbody>
</table>
APPENDIX D. Stimuli Experiment 4 (SAMPA Alphabet)

<table>
<thead>
<tr>
<th></th>
<th>TRAINING</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NV</td>
<td>N_NV</td>
<td>NN_V</td>
</tr>
<tr>
<td></td>
<td>dubo</td>
<td>ti_dubo</td>
<td>duti_bo</td>
</tr>
<tr>
<td></td>
<td>tipa</td>
<td>du_tipa</td>
<td>tidu_pa</td>
</tr>
<tr>
<td></td>
<td>gebo</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>kEpa</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>TEST</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NV</td>
<td>N_NV</td>
<td>NN_V</td>
<td>NV_NV</td>
<td>NVN_V</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Grammatical</td>
<td>Ungrammatical</td>
<td>Grammatical</td>
<td>Ungrammatical</td>
<td>Grammatical</td>
<td>Ungrammatical</td>
<td>Grammatical</td>
<td>Ungrammatical</td>
<td>Grammatical</td>
</tr>
<tr>
<td></td>
<td>dubo</td>
<td>*dupa</td>
<td>ti_dubo</td>
<td>*du_tibo</td>
<td>duti_bo</td>
<td>*duti_pa</td>
<td>dubo_tipa</td>
<td>*dubo_tibo</td>
<td>duboti_bo</td>
</tr>
<tr>
<td></td>
<td>tipa</td>
<td>*tibo</td>
<td>du_tipa</td>
<td>*ti_dupa</td>
<td>tidu_pa</td>
<td>*tidu_bo</td>
<td>tipa_dubo</td>
<td>*tipa_dupa</td>
<td>tipadu_pa</td>
</tr>
<tr>
<td></td>
<td>gebo</td>
<td>*gepa</td>
<td>kE_gebo</td>
<td>*ge_kEbo</td>
<td>kEge_pa</td>
<td>*gekE_pa</td>
<td>gebo_kEpa</td>
<td>*gebo_kEbo</td>
<td>gebokE_bo</td>
</tr>
<tr>
<td></td>
<td>kEpa</td>
<td>*kEbo</td>
<td>ge_kEpa</td>
<td>*kE_gepa</td>
<td>gekE_bo</td>
<td>*kEge_bo</td>
<td>kEpa_gebo</td>
<td>*kEpa_gepa</td>
<td>kEpage_pa</td>
</tr>
</tbody>
</table>
APPENDIX E. Stimuli Experiments 5 and 6 (SAMPA Alphabet)

<table>
<thead>
<tr>
<th>TRAINING</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NV</td>
<td>N_NV</td>
<td>NN_V</td>
</tr>
<tr>
<td>meku dRu</td>
<td>meki_Rabu dRu</td>
<td>meki Rabu_dRu</td>
</tr>
<tr>
<td>meki dRu</td>
<td>meku_Rabi dRu</td>
<td>meku Rabi_dRu</td>
</tr>
<tr>
<td>Rabu dRu</td>
<td>Rabi_meku dRu</td>
<td>Rabi meku_dRu</td>
</tr>
<tr>
<td>Rabi dRu</td>
<td>Rabi_meki dRu</td>
<td>Rabi meki_dRu</td>
</tr>
<tr>
<td>lEpu stu</td>
<td>lEpi sti</td>
<td>govu stu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>govi sti</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TEST</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NV</td>
<td>N_NV</td>
<td>NN_V</td>
<td>NV_NV</td>
<td>NVN_V</td>
<td></td>
</tr>
<tr>
<td>Grammatical</td>
<td>Ungrammatical</td>
<td>Grammatical</td>
<td>Ungrammatical</td>
<td>Grammatical</td>
<td>Ungrammatical</td>
</tr>
<tr>
<td>meku dRu</td>
<td>*meku dRu</td>
<td>meku_Rabu dRu</td>
<td>*meku_Rabi dRu</td>
<td>mekuRabi_dRu</td>
<td>*meku Rabi_dRu</td>
</tr>
<tr>
<td>meki dRu</td>
<td>*meki dRu</td>
<td>meki_Rabu dRu</td>
<td>*meki_Rabi dRu</td>
<td>meki Rabu_dRu</td>
<td>*meki Rabu_dRu</td>
</tr>
<tr>
<td>Rabu dRu</td>
<td>*Rabu dRu</td>
<td>Rabu_meki dRu</td>
<td>*Rabu_meki dRu</td>
<td>Rabu meki_dRu</td>
<td>*Rabu meki_dRu</td>
</tr>
<tr>
<td>Rabi dRu</td>
<td>*Rabi dRu</td>
<td>Rabi_meki dRu</td>
<td>*Rabi_meki dRu</td>
<td>Rabi meki_dRu</td>
<td>*Rabi meki_dRu</td>
</tr>
<tr>
<td>lEpu stu</td>
<td>*lEpu sti</td>
<td>*govi_lEpu sti</td>
<td>govi lEpu sti</td>
<td>*govi lEpu sti</td>
<td>lEpu govi sti</td>
</tr>
<tr>
<td>lEpi sti</td>
<td>*lEpi sti</td>
<td>*govi_lEpi sti</td>
<td>govi lEpi sti</td>
<td>*govi lEpi sti</td>
<td>lEpi govi sti</td>
</tr>
<tr>
<td>govu stu</td>
<td>*govu sti</td>
<td>*lEpi_govu sti</td>
<td>lEpi govu sti</td>
<td>*lEpi govu sti</td>
<td>govu lEpi sti</td>
</tr>
<tr>
<td>govi sti</td>
<td>*govi sti</td>
<td>*lEpu_govi sti</td>
<td>lEpu govi sti</td>
<td>*lEpu govi sti</td>
<td>lEpi govi sti</td>
</tr>
</tbody>
</table>
FRENCH SUMMARY

L'acquisition du langage, et en particulier l'acquisition de la syntaxe, est à première vue l'une des tâches les plus difficiles auxquelles les nourrissons doivent faire face. Les enfants apprennent cependant à parler sans effort apparent et dans un temps limité. Il a été proposé (Chomsky, 1980) que les informations disponibles dans l'input ne sont pas suffisantes pour que les enfants puissent déduire la grammaire correcte (argument de la pauvreté du stimulus). Si cela est le cas, comment les enfants parviennent-ils à la connaissance grammaticale? Deux approches théoriques principales, la théorie de la Grammaire Universelle et les théories basées sur l’usage, essayent de répondre à cette question en étudiant les mécanismes que les apprenants exploitent pour acquérir des connaissances linguistiques. La théorie de la Grammaire Universelle postule que les nourrissons sont dotés de contraintes innées qui leur permettent de choisir la bonne grammaire parmi toutes les grammaires possibles. Selon les théories basées sur l’usage (Tomasello, 2000; Tomasello & Brooks, 1998), les nourrissons ne serait pas équipés des connaissances linguistiques abstraites, mais ils passeraient progressivement d'une connaissance concrète des mots et des expressions aux constructions plus abstraites typiques de la grammaire adulte. Des données empiriques en faveur de l’une ou de l’autre approche ont été récoltées dans le cadre d’expériences utilisant des stimuli du langage naturel ou des paradigmes d’apprentissage de langues artificielles.

Au cours des 50 dernières années, le paradigme de l’apprentissage de langues artificielles a été largement utilisé pour étudier les processus cognitifs impliqués dans l’acquisition du langage. Les langues artificielles imitent les aspects structurels du langage naturel et permettent d’explorer des caractéristiques ciblées de ce dernier (par exemple, la segmentation, l'identification des catégories, etc.). Dans la littérature, certaines études se sont intéressées aux mécanismes utilisés par les nourrissons et les adultes pour effectuer la segmentation lexicale (voir par exemple Aslin et al, 1998; Saffran, Aslin, et al, 1996; Saffran et al, 1999); d'autres études ont examiné les relations syntaxiques d'apprentissage (voir par exemple Reber, 1967). Notre travail s'inscrit dans ce second groupe d'études, car il vise à étudier l'apprentissage de dépendances syntaxiques. Avec ce travail, nous cherchons également à étudier les facteurs phonologiques, prosodiques et sémantiques qui influencent l'apprentissage des règles grammaticales.

Afin de contribuer à l'étude de l'apprentissage des dépendances syntaxiques, nous avons examiné l'apprentissage de deux types de règles dans le cadre de l'apprentissage d'une langue artificielle. La première règle est caractérisée par des dépendances syntaxiques linéaires basées sur la position des éléments dans la séquence linéaire des phrases. La deuxième règle est caractérisée par des dépendances syntaxiques hiérarchiques basées sur la position des éléments dans la structure hiérarchique de la phrase. Dans toutes les expériences de ce travail, les participants effectuaient un entraînement durant lequel ils écOUTaient les phrases caractérisées par la dépendance syntaxique à apprendre. L’entraînement était suivi par des phases de test durant lesquels leur capacité à extraire les régularités syntaxiques était étudiée. Deux objectifs principaux caractèrisent cette thèse: le premier est de déterminer dans quelle mesure différents types de règles basées sur des dépendances syntaxiques linéaires et hiérarchiques peuvent être apprises dans le cadre d’une langue artificielle, le second objectif est d'étudier les facteurs (à savoir, la phonologie, la prosodie, la sémantique) qui influencent l'apprentissage.
L’APPRENTISSAGE DU LANGUAGE

Les enfants apprennent à parler sans effort apparent et dans un temps limité. À 6 mois ils commencent à babiller, c'est à dire qu'ils produisent des séquences de syllabes identiques; à 10 - 12 mois les enfants produisent leurs premiers mots et à 20 - 24 mois ils mettent des mots ensemble dans des phrases simples. À l'âge de quatre ans, les enfants sont capables de maîtriser presque toutes les constructions syntaxiques et possèdent un lexique étendu. Un ensemble de questions auxquelles les chercheurs tentent de répondre concerne le type d'information que les enfants sont capables d'extraire de l'input, comment ils procèdent, et si ils sont équipés avec des contraintes innées. L’argument de la pauvreté du stimulus et sa formalisation ont été attribués à Chomsky (1980). Cependant, comme Pullum et Scholz (2002) l’ont souligné il n'a pas été défini de manière univoque dans la littérature. Par ailleurs, il pourrait être considéré comme une collection d’« affirmations qui se renforcent mutuellement » (Laurence & Margolis, 2001, p. 222) plutôt que comme un argument unique.

Deux caractéristiques principales déterminent l’input auquel les nourrissons sont exposés: 1) l'input ne contient qu'un nombre fini de phrases possibles, 2) il ne contient pas des phrases qui falsifient l'hypothèse erronée (Cowie, 2010). L’input apparait donc insuffisant pour permettre l’apprentissage complet de la grammaire. Malgré cela, les nourrissons réussissent à acquérir le langage et convergent sur la grammaire de leur communauté linguistique. Comment les enfants parviennent-ils à la connaissance grammaticale? Deux principales approches théoriques tentent de répondre à cette question en étudiant les mécanismes que les apprenants utilisent pour acquérir des connaissances linguistiques: la théorie de la grammaire universelle et la théorie basée sur l’usage.

La Grammaire Universelle

Il a été proposé que les informations disponibles dans l'input ne soient pas suffisantes pour que les enfants puissent inférer la grammaire de leur langue. Pour cette raison, Chomsky (1959, 2000) a proposé que les nourrissons soient équipés avec des contraintes qui leur permettent de choisir la bonne grammaire parmi toutes les grammaires possibles. Cette théorie, appelée Grammaire Universelle, a été développée à partir de l'observation que les langues humaines partagent des caractéristiques à un niveau abstrait. La Grammaire Universelle est conçue comme un système de Principes et de Paramètres. Les Principes représentent les propriétés invariantes des langues, tandis que les Paramètres sont des choix binaires qui sous-tendent la variabilité entre les langues. Par exemple, le fait que dans les phrases, les constituants sont organisés suivant l'organisation hiérarchique (Chomsky, 1957) est un principe de la Grammaire Universelle. L'organisation hiérarchique permet la combinaison d'unités de base dans des schémas plus complexes. Cette combinaison permet de générer des phrases grammaticalement complexes et infinies. Comme Rizzi (2007) l’a souligné, l'introduction de la notion de « paramètre » a modifié le point de vue sur l'acquisition des langues. Les apprenants sont censés être pré-équipés avec les Principes et les Paramètres, et leur tâche consiste à définir la valeur de ces derniers en fonction de l’input auquel ils sont exposés. L'apprentissage est considéré comme la sélection des valeurs appropriées des Paramètres parmi toutes les possibilités. Cette sélection est basée sur l'expérience linguistique. Par conséquent, la grammaire d'une langue spécifique est la Grammaire Universelle avec les paramètres...
fixés avec certaines valeurs. En conséquence, afin d'étudier l'acquisition de la syntaxe, il est nécessaire d'explorer la fixation progressive des Paramètres par le nourrisson. Cette approche offre un modèle plausible de l'acquisition de la syntaxe car il introduit une grammaire avec un nombre limité de règles.

**Les théories basées sur l’usage**

Un autre cadre suggère que la connaissance linguistique des nourrissons pourrait être fondée sur l'imitation et des « item-based constructions ». Il a été proposé que les enfants passent progressivement d'une connaissance concrète des mots et des expressions aux constructions abstraites de la grammaire adulte. Par conséquent, les enfants ne seraient pas dotés d'une connaissance abstraite linguistique innée. Au contraire, comme le suggère Tomasello (2000, p. 156) «les enfants apprennent par imitation des expressions linguistiques concrètes par ce qu'ils entendent autour d'eux ».

Cette approche est connue sous le nom de la théorie basée sur l’usage (Tomasello, 2000; Tomasello & Brooks, 1998). La théorie est fondée sur l'approche constructiviste, qui considère la grammaire des adultes comme un inventaire des constructions (Goldberg, 1995). La caractérisation générale de l'acquisition du langage proposée par Tomasello (2003), et clairement résumée par Ambridge (2004), est la suivante:

a) L'enfant utilise l’attention conjointe et l'apprentissage culturel pour apprendre un certain nombre d’énoncés, leur fonction communicative, et leur sens.

b) L'enfant décompose ces énoncés en composantes et généralise à travers les énoncés qui partagent des éléments lexicaux (par exemple, Où est maman? Où est papa? Où est X?).

c) L'enfant détecte les analogies entre ces constructions partiellement abstraites et se déplace vers des constructions plus abstraites (par exemple, SVO construction transitive).

d) L'enfant restreint l'utilisation de certains éléments lexicaux à des constructions particulières dans lesquelles ils sont considérés comme grammaticaux. Par exemple, si un verbe est toujours présenté dans la même structure tel que dormir dans: le garçon dort, le lapin dort, le nounours dort, alors l'enfant en déduit que l'utilisation de ce verbe dans une construction telle que le transitif, n'est pas acceptable.

e) Dans le même temps que a), b), c) et d), l'enfant utilise les informations résultant de l'analyse de l'input pour former les catégories syntaxiques (par exemple, verbe, nom, etc.)

Pour résumer, l’approche basée sur l’usage suggère que les enfants ne peuvent pas être dotés de connaissances linguistiques abstraites et innées. Au lieu de cela, pendant les premières années de la vie ils passeraient progressivement des constructions liées à des objets à des constructions abstraites. Initialement les enfants comptent sur l'imitation des constructions plus concrètes et
seulement dans une phase ultérieure du développement, ils développent plusieurs schémas abstraits (par exemple, SVO construction transitive) sur la base de leur capacité à repérer des similitudes.

L’analyse syntaxique

L’analyse syntaxique permet la compréhension de phrases en décrivant comment les mots sont structurés dans une représentation hiérarchique en fonction de la grammaire d'une langue. L’interprétation d'une phrase est en effet déterminée par la façon dont les mots sont combinés les uns avec les autres. Comme montré dans l'exemple (6), une phrase en Anglais peut avoir deux significations différentes selon la façon dont les mots sont hiérarchiquement structurés.

(6) The girl hit the man with the umbrella.

The girl [[hit [the man]] [with the umbrella]].

The girl [hit [the man [with the umbrella]]].

Selon une première interprétation, le parapluie est utilisé comme un instrument. Selon la seconde interprétation, le parapluie est une propriété de l'homme. Boland et Blodgett (2001) fournissent une schématisation claire de l'analyse syntaxique, qui est décrite comme constituée de trois composantes:

1. Génération de la structure syntaxique (ou des structures alternatives multiples, voir les théories Interactives)
2. Sélection d'une structure syntaxique (si plus d'une)
3. Ré-analyse si la structure choisie est incorrecte

Ces trois composantes ne sont pas nécessairement mises en œuvre à différents stades dans toutes les théories d'analyse. Dans les théories d'analyse qui supposent qu’une seule structure syntaxique est générée, comme dans le modèle « garden-path » (Frazier, 1987a), il n'y a pas de distinction entre les deux premières composantes. Habituellement, ce processus est réalisé rapidement et de façon précise, mais parfois les gens peuvent éprouver quelques difficultés. C’est le cas lorsque nous traitons des phrases ambiguës.

Les modèles modulaires - Ces modèles sont basés sur l’hypothèse que l'esprit est composé de différents modules (Fodor, 1983), qui réalisent des processus distincts et spécifiques. L'idée générale partagée par ces modèles est que le processeur du langage est composé de plusieurs sous-systèmes qui transmettent les informations selon une direction spécifique (traitement en série). Ces modèles sont dits modulaires car le processeur est supposé exploiter principalement (ou uniquement) des informations syntaxiques dans les étapes initiales de l'analyse. Seule une analyse à la fois est adoptée par le processeur et une seule et même structure syntaxique est donc activée à la fois. Le modèle le plus connu est le modèle « garden-path » proposé par Frazier et ses collaborateurs (Frazier, 1987a).
Les modèles interactifs - Ces modèles sont basés sur l'hypothèse que des analyses multiples sont effectuées par le processeur et, par conséquent, plus d'une structure syntaxique est activée à la fois; toutes les sources d'information pertinentes peuvent être utilisées immédiatement au cours du traitement de la phrase (MacDonald, 1993 ; McClelland, John & Taraban, 1989). L'analyse qui est soutenue par divers types d'informations (par exemple, syntaxique, sémantique, etc.) est la plus active. Les difficultés rencontrées dans le traitement des phrases ambiguës sont ici attribuées à la cooccurrence de plusieurs alternatives. Lorsque deux alternatives ont un niveau d'activation similaire, le traitement est plus complexe.

Apprentissage des Langues Artificielles

Afin d'acquérir plus d'informations sur les mécanismes qui permettent l'apprentissage des langues naturelles, les chercheurs se sont penchés sur l'apprentissage des langues artificielles (ALL). Dans tous ces paradigmes, les participants sont exposés à du matériel structuré selon certaines règles. Ces règles imitent les aspects structurels de la langue naturelle, mais sont simplifiées par rapport aux conditions réelles d'apprentissage des langues. Ces stimuli sont contrôlés et présentés aux participants lors d'une phase d'entraînement, par la suite les participants sont testés afin d'étudier leur capacité à extraire et à généraliser ces régularités. Il s'agit donc d'utiliser des systèmes simplifiés pour investiguer les différents aspects du traitement linguistique indépendamment les uns des autres. Initialement, comme deVries (2011) le fait remarquer, les paradigmes ALL ont été utilisés pour étudier l'apprentissage implicite. Par la suite, les paradigmes ALL ont été employés pour étudier les différents aspects de l'apprentissage du langage naturel de manière isolée. Gomez et Gerken (2000) ont identifié quatre sujets d'étude principaux: la segmentation en mots, l'apprentissage de l'ordre des mots, l'apprentissage des relations grammaticales, et l'apprentissage des catégories syntaxiques.

Une des premières difficultés à laquelle les nourrissons doivent faire face dans l'apprentissage des langues est de savoir comment segmenter les mots dans le flux de parole continue qu’ils entendent. Des études ont montré que les nourrissons possèdent un mécanisme sophistiqué qui leur permet de segmenter le discours sur la base de calculs statistiques (par exemple, Saffran, Aslin, & Newport, 1996; Saffran, Newport, & Aslin, 1996). Ces études fondamentales ont été suivies par d'autres montrant que ce mécanisme d'apprentissage statistique peut être exploité pour identifier des unités dans d’autres domaines tels que le domaine visuel ou le domaine musicale (par exemple, Fiser & Aslin, 2002a; Gebhart, Newport, & Aslin, 2009; Saffran, Johnson, Aslin, & Newport, 1999).

Les bébés doivent non seulement apprendre à segmenter les mots, ils apprennent également comment les mettre dans le bon ordre dans les phrases. Les études ALL ont été utilisées pour déterminer si les bébés étaient capables de discriminer des phrases générées par différentes grammaires basées sur un ordre des mots différent. Ces études ont suggéré que les nourrissons sont capables de le faire d'une façon rapide et précise (Gomez & Gerken, 1999). D'autres recherches ont montré que les bébés ne sont pas seulement capables d'acquérir l'ordre des mots dans les phrases mais qu'ils sont aussi capables d'extraire des patterns plus abstraits. Très tôt dans leur vie, ils sont
capables de généraliser leur connaissance des motifs abstraits à de nouvelles situations (par exemple, un nouveau lexique) (Marcus, Vijayan, Bandi Rao, & Vishton, 1999).

Enfin, le dernier thème exploré dans les expériences ALL est l'apprentissage des catégories syntaxiques. La capacité d'abstraction des catégories est très importante pour l'acquisition du langage. En fait, si nous reconnaissions qu'un nouveau mot appartient à une certaine catégorie syntaxique nous savons automatiquement que toutes les propriétés qui sont valides pour cette catégorie s'appliquent au nouveau mot (par exemple, si un anglophone reconnaît que WUG est un nom, il devrait être capable de créer le pluriel WUGS, même si il n'a jamais entendu cette expression). Ce type d'abstraction a été étudié avec les paradigmes ALL dans lesquels les participants ont appris à classer les nouveaux mots dans des catégories et ensuite ils ont appris les relations entre ces catégories. Plusieurs expériences ont montré que les adultes sont capables de catégoriser les nouveaux mots sur la base d'indices perceptifs qui indiquent l’appartenance à une catégorie (par exemple, Frigo & McDonald, 1998).

Il y a plusieurs avantages à utiliser les paradigmes ALL pour l’étude de l'apprentissage des langues. Gomez et Gerken (2000) en ont souligné deux: 1) les études ALL permettent à l’expérimentateur de contrôler l’input auquel les participants sont exposés; 2) les études ALL permettent à l’expérimentateur de contrôler l’apprentissage préalable. Le contrôle de l’input que les apprenants reçoivent a permis l’étude des aspects de la langue de façon isolée. Puisque le langage naturel est un système complexe, il est difficile de trier les divers aspects qui jouent un rôle lors de l'acquisition. Par conséquent, l'utilisation de langages artificiels permet aux chercheurs d'éviter les caractéristiques redondantes et / ou les confusions qui peuvent influencer le processus d'apprentissage. Malgré ces avantages, les paradigmes ALL doivent toujours être confrontés à des conditions réaliste d'acquisition du langage afin d'améliorer leur validité.

QUESTIONS DE RECHERCHE

possibilité d'apprendre différents types de règles basées sur des dépendances syntaxiques linéaire et hiérarchique dans une langue artificielle, le deuxième objectif était d'étudier les facteurs influençant l'acquisition des règles. Ces questions peuvent être formulées comme suit:

Q1. Les adultes sont-ils capables d'apprendre des dépendances syntaxiques dans une langue artificielle

   a) Quand elles sont caractérisées comme des dépendances linéaires?

   b) Quand elles sont caractérisées comme des dépendances hiérarchiques?

Q2. Quels sont les facteurs qui influencent l'apprentissage des règles dans une langue artificielle?

   Trois facteurs, et leur influence dans l'apprentissage des règles, ont été prises en compte dans cette thèse et ont été étudiés dans six expériences. Premièrement, nous avons étudié si le type d'unités utilisées dans le langage artificiel (linguistique vs non-linguistiques) influence l'apprentissage de la règle de l'accord. Deuxièmement, nous avons étudié si des indices phonologiques et prosodiques dans le signal influencent la création de relations entre les éléments impliqués dans la règle syntaxique. Cette enquête nous a permis d'explorer si, pour apprendre une règle syntaxique complexe, la cooccurrence de plusieurs indices est nécessaire. Enfin, nous avons exploré l'influence des indices sémantiques dans l'apprentissage d'une langue artificielle.

   Par rapport aux études antérieures dans le domaine de l'apprentissage des langues artificielles, nous avons introduit trois nouveautés principales. Premièrement, les règles que nous avons étudiées étaient plus proches de celles que nous trouvons dans les langues naturelles que les règles utilisées dans les études ALL précédentes. Deuxièmement, nous avons abordé la question du rôle des indices sémantiques dans une étude d'apprentissage des langues artificielles. Troisièmement, nous avons étudié différents types de connaissances et de généralisations.

**CADRE EMPIRIQUE**

Afin d'étudier l'apprentissage de structures syntaxiques complexes, nous avons créé deux règles d'accord sujet-verbe. Nous avons décidé d'utiliser l'accord sujet-verbe car il est considéré comme un exemple prototypique de la dépendance à la structure. En effet, dans les langues naturelles, le sujet d'une phrase s'accorde avec son verbe, même les deux mots ne sont pas contigus dans la séquence linéaire des mots. Le sujet et le verbe sont liés en vertu de leur position dans la configuration hiérarchique de la phrase. Par conséquent, l'accord est un phénomène qui est calculé sur la base des relations hiérarchiques (par exemple c-commande, localité) entre les constituants de la phrase.

Dans cette thèse nous avons créé deux règles d’accord. La règle utilisée dans les expériences 1 et 2 est basée sur la position des mots dans la séquence linéaire de la phrase, alors que la règle utilisée dans les expériences 3, 4, 5 et 6 est fondé sur l'organisation hiérarchique des constituants de la phrase.
Règle linéaire d'accord sujet-verbe (Expériences 1 et 2)

La règle linéaire d'accord sujet-verbe a été définie comme suit: le verbe (V) s'accorde avec le premier nom (N) dans la phrase. Cette règle d'accord était basée sur la position linéaire des éléments dans une phrase. Les mots ont été combinés pour constituer trois différents types de structures de phrase: NV, NVN, NNV. En raison de la structure choisie, deux types de dépendances syntaxiques ont été étudiés: a) une dépendance adjacente lorsque le verbe et le sujet étaient contigus dans les structures, comme dans NV et NVN; b) une dépendance non-adjacente où il y avait un intervenant entre le sujet et le verbe, comme dans la structure NNV.

Règle hiérarchique d'accord sujet-verbe (Expériences 3, 4, 5, 6)

La règle hiérarchique d'accord sujet-verbe a été définie comme suit: le verbe (V) s'accorde avec le nom (N) le plus haut dans son constituant; s'il n'y a pas de nom dans ce constituant, le verbe s'accorde avec le nom le plus haut dans le constituant qui précède. Selon cette règle, la relation d'accord entre le sujet et le verbe est déterminée par la structure des constituants de la phrase. Par ailleurs, dans nos expériences un indice prosodique (i.e., une pause) a été utilisé pour marquer les limites des constituants. Trois différents types de structures de phrases ont été créés:

a) Phrases de deux mots se composant d'un nom et d'un verbe, NV; dans ce cas, la règle de l'accord détermine une dépendance adjacente.

b) Phrases de trois mots se composant de deux noms et d'un verbe; dans des phrases NN_V (où "_" indique la position de la pause) la règle d'accord détermine une dépendance non-adjacente. Dans les phrases N_NV la règle d'accord détermine une dépendance adjacente.

c) Phrases de quatre mots se composant de deux noms et de deux verbes; dans des phrases NV_NV la règle d'accord détermine deux dépendances adjacentes; dans les phrases NVN_V la règle d'accord détermine à la fois une dépendance adjacente et une dépendance non-adjacente, les deux verbes s'accordent avec le sujet qui était situé au début de la phrase.

ANALYSE DES DONNÉES

L'hypothèse générale qui guide l'interprétation des résultats de paradigmes d'apprentissage des langues artificielles est que si les participants ont des scores nettement au-dessus de la chance ils ont appris la grammaire (ou la règle) de l'input. La mesure la plus largement utilisée pour ce genre d'études est la précision de réponse, définie comme le pourcentage de réponses correctes sur le nombre total d'items. Cette mesure est ensuite comparée au niveau de chance (par exemple, le 50% pour une tâche avec deux réponses possibles, comme c'est le cas pour la tâche de jugement de grammaticalité). Dans cette thèse, nous avons utilisé une autre mesure (i.e., le d-prime) qui n'est pas souvent prise en compte dans les études d'apprentissage des langues artificielles, mais peut néanmoins représenter un bon moyen de caractériser les performances des participants. Dans le cadre de nos expériences, il représente un indice de la capacité à discriminer entre les éléments grammaicaux et agrammaticaux. En d'autres termes, cette mesure représente un indice de la
sensibilité des participants à la grammaticalité. L'utilisation de cette mesure permet de contourner le problème concernant les biais des participants.

Dans les expériences d'apprentissage des règles, il est possible que les participants utilisent des règles alternatives pour accomplir la tâche. Dans ce cas, leurs réponses ne reflètent pas la représentation de la règle cible. L'exploration de ces règles alternatives peut conduire à des interprétations différentes des résultats. L'hétérogénéité des résultats, si correctement prise en considération, peut fournir de nouvelles informations quant aux processus impliqués dans l'apprentissage des langues artificielles. Comment pourrions-nous examiner les règles que nos participants ont utilisées pendant la tâche de jugement de grammaticalité? Afin d'étudier quelle règle les participants ont utilisée, nous avons effectué une analyse des « templates ». Nous avons défini a priori un ensemble de règles alternatives plausibles étant donné l'input lors de l'entraînement, que nous avons appelé «templates». Le but de l'analyse des « templates » a été de trouver un template qui caractérise le mieux le comportement des participants. Pour atteindre cet objectif, nous avons analysé chaque ensemble de données compte tenu de la forme de plusieurs « templates » possibles pour chaque expérience. Le « template » avec la plus haute valeur d-prime a été considéré comme le plus représentatif du comportement du participant.

TYPES DE CONNAISSANCE

Toutes les expériences comprenaient des phases d'entraînement et de test. Dans l'entraînement, les participants devaient apprendre des phrases générées par la règle d'intérêt, et dans la phase de test, ils devaient effectuer une tâche de jugement de grammaticalité dans laquelle ils devaient déterminer si les phrases présentées avaient été générés par la règle ou non. Comme différents types de matériel de test ont été utilisés, nous avons été en mesure d'étudier différents types de connaissances, à savoir, la mémorisation (quand les phrases grammaticales étaient les mêmes que dans la phase d'entraînement), la généralisation lexicale (quand les phrases grammaticales avaient un lexique différent de celles de la phase d'entraînement, mais les mêmes structures syntaxiques), la généralisation syntaxique (lorsque les phrases avaient différentes structures syntaxiques par rapport à celles de la phase d'entraînement, mais le même lexique), et la généralisation lexicale et syntaxique (lorsque les phrases grammaticales avaient différentes structures syntaxiques et un lexique différent par rapport à celles de la phase d'entraînement).

EXPÉRIENCES 1 ET 2

Dans les deux premières expériences présentées dans cette thèse nous avons étudié si les adultes étaient capables d'apprendre une règle linéaire d'accord sujet-verbe dans deux domaines différents, la musique et le langage. Pour cela, les deux études mesuraient la capacité d'apprendre une règle (i.e., le verbe s'accorde avec le premier nom dans la phrase) fondée sur la position des éléments dans la chaîne (c'est à dire, la phrase). Le déroulement était presque identique pour les deux études. Les expériences comprenaient une phase de familiarisation au Lexique, une phase d'entraînement, et une phase de test. Dans la phase de familiarisation au Lexique, les participants étaient familiarisés avec les noms et les verbes du lexique à travers l'écoute passive de tous les mots (mots-tons ou pseudo-mots). Les phases d'entraînement et de test constituaient la session.
d'apprentissage de la règle. Dans la phase d’entraînement, les participants étaient invités à écouter un échantillon de phrases suivant la règle linéaire d’accord. Dans la première expérience les participants devaient effectuer une tâche de détection à chaque fois qu’ils écoutaient un mot pluriel alors que dans la seconde expérience, ils écoutaient passivement les phrases. Dans la phase de test, les participants devaient effectuer une tâche de jugement de grammaticalité.

Dans la première expérience le vocabulaire était constitué de ton-mots, chaque mot était composé de trois notes de piano. Une hauteur différente déterminait la distinction entre les catégories grammaticales, aigu pour les noms et grave pour les verbes. La distinction entre le singulier et le pluriel des mots-ton était mise en œuvre par l'ajout d'un suffixe, à savoir une courte note (165 ms G4) à la fin des mots au pluriel. Le même ton a été utilisé pour les noms au pluriel et les verbes au pluriel. En conséquence, les mots au singulier étaient constitués de trois tons (durée totale = 990 ms) et les mots pluriels étaient constitués de quatre tons (durée totale = 1155 ms). Les phrases agrammaticales étaient caractérisées seulement par la violation de la règle accord (i.e., les phrases avec un sujet singulier et un verbe au pluriel ou des phrases avec un sujet pluriel et un verbe au singulier). Aucune violation de position a été introduite (par exemple, il n'y avait pas des phrases où le verbe était en première position). Si le sujet et le verbe étaient adjacents dans la structure linéaire de la phrase, comme dans la structure de NVN, la phrase était dans l'état adjacent; s'il y avait un élément intervenant entre le sujet et le verbe, comme dans la structure de NNV, la phrase appartenait à la condition non-adjacente. Les noms dans les phrases de trois mots, avaient soit le même nombre, auquel cas ils étaient dans la condition « matching » (en français, appariée) soit un nombre différent et donc appartaient à la condition « mismatching » (en français, non-appariée).

Nous avons réalisé l'expérience 2 avec des stimuli linguistiques, ayant pour objectif d'étudier la différence entre l'apprentissage des régularités dans un domaine non-linguistique et dans un domaine linguistique. Les mots impliqués dans cette règle étaient trisyllabiques. Ils étaient combinés pour former des phrases à deux ou trois mots. La distinction entre les catégories grammaticales était déterminée par l'utilisation de consonnes avec différents points d’articulation, à savoir les consonnes occlusives pour les noms et les consonnes fricatives pour le verbe. La distinction entre le singulier et pluriel des mots était implémentée par l’utilisation de différentes voyelles finales.

Dans l’Expérience 1, nous avons observé que les participants ont répondu au hasard dans la tâche de jugement de grammaticalité des phrases NV (t (17) = 1.293, p = 0.106). Ce résultat est particulièrement surprenant, car ces phrases étaient censées être les plus faciles. Pour répondre correctement à ces phrases, les participants devait savoir seulement que l'accord était marqué par la présence simultanée (ou l'absence) d'un ton final identique sur les deux mots-ton. Cette mauvaise performance suggère que les participants ont été incapables d'utiliser cette information sur le nombre pour juger la grammaticalité de la phrase et, par conséquent, ils n'ont pas appris la règle.

Deuxièmement, nous avons observé un effet principal de la variable Match (F (1,17) = 13.298, p = 0.002). Cet effet révèle que la performance des participants était sensiblement meilleure pour les phrases de la condition matching (M = 0.67, SD = 0.99) que pour les phrases dans la condition mismatching (M = -0.08, ET = 0.79). Dans les phrases grammaticales avec trois mots,
tous les mots avaient le même nombre et donc le même ton final. En conséquence, les bonnes performances observées dans les phrases matching ne démontrent pas si les participants ont appris la façon dont les catégories grammaticales étaient marquées et / ou le rôle syntaxique «sujet» attribué.

Troisièmement, dans l'analyse de « templates », nous avons observé que quatre participants ont appris le modèle 1^2^3^ (i.e., tous les mots ont la même terminaison), tandis que le reste des participants n'ont appris aucun « template ». Nous avons étudié si on pouvait reconduire l'effet de la variable Match à la performance de ces quatre participants, comme il était possible que la meilleure performance observée dans les phrases matching soit entièrement attribuable à la performance de ces quatre participants. Nous avons effectué une analyse de variance ne considérant que les 14 participants qui n'ont pas appliqué la template 1^2^3^ pour examiner si l'effet du matching était encore significatif ou non. Nous avons observé que l'effet principal de la variable Match était encore significatif (F (1,13) = 6.341, p = 0.026). Ainsi, même les participants qui n'ont pas appris ce template, jugeaient les phrases matching mieux que les phrases mismatching. Comment peut-on expliquer ce fait?

Comme mentionné précédemment, les phrases matching dans la condition grammaticale pour NNV et NVN étaient perceptivement saillantes, puisque tous les trois mots dans ces conditions avaient le même nombre. Il est possible que même les participants qui n'ont pas appliqué systématiquement le template 1^2^3^ ont bénéficié de cette caractéristique de la phrase. Cependant, si tel était le cas, nous devrions avoir observé plusieurs jugements corrects non seulement pour des phrases grammaticales, mais aussi pour les phrases agrammaticales. Les participants auraient dû rejeter les phrases agrammaticales facilement, puisque le verbe avait un nombre différent comparé à celui des deux noms. Puisque d'-prime ne nous permet pas de différencier entre phrases grammaticales et non grammaticales, nous avons analysé la précision de réponse. Les performances au-dessus de la chance (70% de précision) étaient limitées aux phrases grammaticales; les phrases agrammaticales étaient jugées au hasard (46% de précision). Par conséquent, l'explication selon laquelle les participants ont basé leurs jugements sur la saillance perceptive des phrases matching ne peut pas rendre compte de la performance pour les phrases agrammaticales, qui aurait été systématiquement rejetées, mais l'étaient pas. Le fait que les phrases grammaticales ont été jugées meilleures que les phrases agrammaticales pourrait être dû à la mémorisation, car seul les premières, mais pas les secondes, ont été présentées lors de la phase d’entraînement.

Compte tenu de la faible performance des participants sur les phrases NV et de leurs valeurs d'-prime systématiquement faibles sur les autres types de phrases, nous ne pouvons que conclure que nos participants n'ont pas appris la règle. Cependant, il est plus difficile de déterminer ce qu’ils ont appris et comment ils ont géré leurs réponses dans la tâche de jugement de grammaticalité. Il y a deux scénarios possibles. Premièrement, les participants ont pu être biaisés en faveur de la réponse "grammatical". Cela pourrait expliquer les faibles valeurs d'-prime, car cette mesure prend en compte les biais des participants, et aussi la meilleure performance sur les phrases grammaticales matching par rapport aux phrases agrammaticales matching. Afin de tester cette hypothèse nous avons considéré le nombre de réponses «grammatical» et «agrammatical » indépendamment de
savoir si la réponse était correcte ou non. Nous avons observé que la réponse «grammatical» représentait les 58% du total des réponses pour les phrases avec trois mots, révélant que la réponse grammatical a tendance à être préférée à la réponse «agrammatical».

Un autre scénario possible pourrait être que les participants ne cesssaient pas de changer leurs stratégies de réponse lors de la phase test, en essayant d'appliquer des règles différentes sans jamais apprendre la bonne. Ce scénario est cependant difficile à examiner en raison du nombre élevé de combinaisons de stratégies que les participants auraient pu adopter. Dans l'ensemble, les résultats actuels suggèrent que les participants ne sont ni en mesure d'apprendre la règle linéaire d’accord, ni d’extraire les informations sur les catégories grammaticales et le nombre. Quatre explications de ce manque d'apprentissage sont possibles:

1) L'utilisation de la tâche de détection lors de la formation aurait empêché l'apprentissage. Bien qu'elle force les participants à prêter attention aux phrases à travers toute la session, la tâche de détection pourrait avoir dirigé leur intérêt sur un seul mot dans les phrases, les empêchant ainsi d’apprendre les relations entre les mots. Cette explication est purement spéculative, car les données que nous possédons ne nous permettent pas de la tester.

2) Les indices utilisés pour marquer les catégories grammaticales et le nombre ne sont pas suffisamment saillants. Toutefois, cette explication irait à l'encontre des conclusions antérieures montrant que les adultes sont capables d'exploiter la différence de hauteur (Creel et al., 2004).

3) Garder la trace des relations entre les tons et entre les mots était trop difficile, car les participants n'ont pas été formés musicalement.

4) La règle linéaire d’accord utilisée dans la présente expérience ne s'apprend pas dans le domaine de la musique. Afin de tester cette explication possible, dans l'expérience suivante, nous avons étudié l'apprentissage de la même règle dans le domaine linguistique, en essayant de vérifier si les performances des participants s’améliorent en changeant le type de stimuli à apprendre.

Dans l’Expérience 2, les performances des participants à des phrases NV a été nettement au-dessus de la chance (t (20) = 5.241, p < .001). Cette observation suggère que les participants ont au moins appris comment l’accord était marqué (i.e., les mots avec le même nombre doivent partager la même voyelle finale). Comme il n'y avait que deux mots dans la phrase, il n'était pas nécessaire de distinguer les catégories grammaticales ou de prendre en compte la position du sujet et du verbe dans la séquence linéaire des mots.

Nous avons observé deux autres résultats importants 1) un effet principal de la variable Match (F (1,20) = 6.033, p = 0.023) ; 2) le fait qu’un grand nombre de participants (i.e., 9 des 21 participants) ont utilisé systématiquement le template 1^2^3^. Encore une fois, ces deux résultats peuvent être reliés, l’effet du Match pouvant être attribué à la performance des participants utilisant ce template. Pour vérifier cette hypothèse nous avons réalisé une ANOVA sur les valeurs d-prime. La variable entre les sujets était le groupe, à savoir si un participant appartient au groupe utilisant le 1^2^3^ template ou non. Les variables intra-sujets étaient Localité et Match. L'analyse a révélé un
effet principal de la variable Match (F (1,19) = 27.172, p <.001): les participants ont réalisé un meilleur score dans les phrases matching (M = 1.33, ET = 1.91) que dans les phrases mismatching (M = 0.24, ET = 1.22). Par ailleurs il y avait une ‘effet principal de la variable Groupe (F (1,19) = 27.172, p <.001). Les participants appartenant au groupe utilisant le template a obtenu de meilleurs résultats (M = 1.67, ET = 1.96) que le reste des participants (M = -0.09, ET = 0.79). L’effet principale de la localité n’était pas significatif (F < 1). Fait plus intéressant encore, l’interaction entre les variables match et groupe est significatif (F (1,19) = 42.906, p <.001). Cette interaction reflète le fait que la différence entre la matching et mismatching n’était significatif que pour le groupe utilisant le template (t (8) = 7.341, p <0.001) et non pour le reste des participants (t (11) = -1.071, p = 0.307). Aucune des autres interactions n’est significative (Fs <1). Nous pouvons conclure que l’effet du match observé était dû aux participants qui ont appliqué le template. Par conséquent, il semble possible d’identifier deux groupes: un premier avec les neuf participants qui ont appris le template 1^2^3^, et un second avec les douze participants qui n’ont pas appris la règle d’accord ou l’un des template alternatifs. Les participants appartenant à ce dernier groupe ont des scores au niveau du hasard pour tous les types de structure de phrase. Parconséquent, nous supposons qu’ils n’ont pas appliqué une règle ou appris tous les aspects de la langue artificielle.

Concernant les participants qui ont appris l’autre template, nous pouvons nous demander pourquoi ils ont appris une règle différente. Il est possible que lors de l’entraînement ils se sont concentrés uniquement sur les phrases grammaticales matching, car elles sont perceptivement plus saillantes que les phrases grammaticales mismatching. En fait, les phrases matching NVN et les phrases matching NNV ont la même séquence de voyelles finales. Ce template peut vraisemblablement être perçu facilement sur la seule base de ces rimes. La présence d’un tel modèle pourrait avoir empêché les participants d’apprendre la règle d’accord correcte et faciliter l’apprentissage de l’autre template incorrect.

Pour résumer, les résultats de l’Expérience 2 suggèrent qu’aucun participant n’a appris la règle linéaire d’accord. Les participants peuvent être divisés en deux groupes selon leurs performances. Un groupe a une performance au niveau du hasard dans toutes les conditions et n’a pas appris la règle linéaire d’accord ou un des template alternatifs. Le deuxième groupe de participants a appris le template 1^2^3^ sur la base de la saillance perceptuelle des phrases grammaticales matching, mais n'a pas appris la règle linéaire d’accord. Les résultats des deux expériences ont révélé que les participants n’ont pas appris la règle dans les deux domaines. Dans les deux expériences il y avait un effet de la variable Match. Toutefois, les analyses proposent différentes origines de cet effet. Nous avons supposé que dans l’Expérience 1, les participants avaient une préférence pour la réponse «grammatical», ce qui expliquerait les faibles valeurs d-prime et les performances au hasard, ou qu’ils avaient changé continuellement leur stratégie de réponse. Dans l’Expérience 2, nous avons distingué deux groupes de participants, un groupe applique le template 1^2^3^ et un groupe répond au niveau du hasard pour toutes les phrases de trois mots. Comme les participants n’ont pas appris la règle dans les deux expériences, nous ne pouvons pas tirer de conclusions définitives au sujet de la comparaison des domaines de la musique et linguistique.
EXPÉRIENCES 3, 4, ET 5

Nous présentons ici trois expériences dont le but est d'étudier l'apprentissage d'une règle hiérarchique d'accord sujet-verbe. La règle hiérarchique d’accord a été définie comme une relation entre des éléments appartenant à différentes catégories grammaticales, à savoir les noms et les verbes, qui occupent des positions spécifiques dans la structure hiérarchique de la phrase, comme par exemple sujet et verbe. Pour apprendre la règle, trois étapes sont nécessaires:

1) Apprendre les catégories grammaticales et le nombre (i.e., que les mots se répartissent en différentes catégories grammaticales et qu’ils peuvent être soit singuliers ou pluriels).

2) Apprendre les marqueurs morphologiques et leur association (par exemple, la cooccurrence des marqueurs).

3) L'apprentissage de la règle hiérarchique d'accord, c’est à dire à analyser la phrase en constituants, afin d'identifier les rôles syntaxiques (i.e., le sujet et le verbe), et à comprendre l'accord entre le sujet et le verbe (par exemple, que dans la phrase N_NV, le sujet est le nom dans le même constituant que le verbe, tandis que dans la phrase NN_V, le sujet est le nom dans le constituant qui précède celui du verbe).

Dans les expériences 3, 4 et 5, les informations sur les catégories grammaticales et le nombre ont été véhiculées par différents indices phonologiques (i.e., caractéristiques des voyelles et des consonnes, le nombre de syllabes et la structure syllabique), tandis que l’information sur la structure en constituants a été fournie par les indices prosodiques (i.e., contour mélodique et les pauses). Plus précisément, dans les Expériences 3 et 4, les catégories grammaticales et le nombre ont été marquées par des indices phonologiques (par exemple, les caractéristiques de consonnes et de voyelles). Ainsi, le point d'articulation des consonnes (i.e., bilabiale vs. alvéolaire) dans la seconde syllabe a été utilisé pour marquer la catégorie grammaticale et le voisement (i.e., voisé vs. non voisé) indiquait le nombre. Les catégories grammaticales ont également été marquées par le degré d'ouverture de la bouche des voyelles dans la seconde syllabe, et le nombre a été marqué par les voyelles finales.

Expérience 3

Dans l'Expérience 3, nous avons étudié l'apprentissage d'une règle hiérarchique d'accord où la relation d’accord entre le sujet et le verbe reposait sur la structure en constituants de la phrase et non pas sur la position linéaire des mots. Le lexique a été constitué par des pseudo-mots bisyllabiques. Les catégories grammaticales et le nombre ont été marqués par des indices phonologiques (par exemple, les caractéristiques de consonnes et de voyelles). Ainsi, le point d'articulation des consonnes (i.e., bilabiale vs. alvéolaire) dans la seconde syllabe a été utilisé pour marquer la catégorie grammaticale et le voisement (i.e., voisé vs. non voisé) indiquait le nombre. Les catégories grammaticales ont également été marquées par le degré d'ouverture de la bouche des voyelles dans la seconde syllabe, et le nombre a été marqué par la profondeur du point d'articulation des voyelles. En croisant tous les indices phonologiques, nous avons obtenu quatre secondes syllabes possibles (i.e.,-du,-bo,-ti,-pa). Des indices prosodiques (i.e., les pauses) ont été employés pour marquer les constituants.
La première syllabe des mots appartenant à un grand ensemble (i.e., 92 syllabes différentes), en ligne avec la grande variabilité des racines des mots dans les langues naturelles. Le rôle de la première syllabe consistait à créer une variabilité dans le signal contrairement à la seconde syllabe, partie la moins variable du signal impliquée dans la règle. La forte variabilité de la racine visait ainsi à concentrer l'attention des participants sur la seconde syllabe. Comme Gomez (2002) l’a observé, quand les adultes et les nourrissons effectuent une tâche dans laquelle ils ont à apprendre les dépendances syntaxiques du type AXC, ils ont tendance à se concentrer principalement sur les dépendances adjacentes. Au contraire, s'il y a une forte variabilité de l'élément X (i.e., 16 à 24 X différents), les participants sont capables d'extraire les dépendances non adjacentes entre les éléments A et C. Contrairement au matériel utilisé par Gomez, dans la présente expérience, l'élément très variable se situe au sein des mots et non entre les mots impliqués dans la relation. Cependant, le rôle de la variabilité reste le même, à savoir mettre en évidence les éléments récurrents impliqués dans la dépendance syntaxique. Nous avons supposé que, puisque les participants se concentreront sur les relations entre les secondes syllabes (c'est à dire, les suffixes), ils seront capables d'exploiter les indices phonologiques qui caractérisent la catégorie grammaticale et le nombre. Une fois que les participants sont capables de détecter les indices phonologiques (par exemple, les marqueurs morphologiques accord) et leur association, ils possèdent les informations nécessaires pour apprendre la règle. Toutefois, rappelons ue pour apprendre la règle hiérarchique d'accord les participants doivent analyser les phrases en constituants en exploitant la présence des pauses. Ils doivent apprendre les rôles syntaxiques et identifier la position du sujet, qui varie en fonction de la structure hiérarchique de la phrase plutôt que de l'ordre linéaire des mots.

Les effets de trois variables caractérisant le matériel ont été étudiés, à savoir la localité, le match et le type de connaissances. Avec le terme « localité » nous nous référerons au fait que la règle hiérarchique d'accord conduit à deux types de dépendances, les dépendances adjacentes quand le sujet et le verbe appartiennent au même constituant (comme dans N_NV), et les dépendances non adjacentes quand ils appartiennent à différents constituants (comme dans NN_V). La variable « match » réfère au fait que les deux noms dans les phrases avaient soit le même nombre soit un nombre différent. Enfin, la variable « type de connaissances » se réfère aux types de connaissances étudiés dans l’expérience en raison des différences et des similitudes entre les caractéristiques du matériel de test et d’entraînement. Dans la présente expérience deux types de connaissances ont été étudiés: la généralisation lexicale et la généralisation lexicale et syntaxique. Le premier a été étudié lorsque le vocabulaire utilisé dans la phase de test était différent de celui utilisé dans la phase d’entraînement, le dernier était présent lorsque le vocabulaire et les structures syntaxiques présentées dans les phases de test étaient différents de ceux utilisés dans la phase d’entraînement.

La procédure employée était la suivante. Dans la phase de familiarisation au lexique, les participants ont été familiarisés au lexique par une présentation auditive des mots ; dans la phase d'apprentissage de la règle, les participants ont été invités à écouter passivement des phrases suivant la règle hiérarchique d'accord pour effectuer ensuite une tâche de jugement de grammaticalité. Nous considérerons ici les deux premières étapes du processus d'apprentissage décrit dans l’introduction à ces expériences. Nous pouvons tester si les participants atteignent ces étapes analysant leurs performances aux phrases matching. Notre hypothèse est que, afin de juger ces phrases, les participants doivent seulement identifier les catégories grammaticales, le nombre et l'accord
morphologique. Il n'était pas nécessaire d'être capable d'analyser les phrases en constituants et d'identifier les rôles syntaxiques. Par conséquent, pour effectuer correctement la tâche de jugement de grammaticalité, les participants n'ont pas besoin d'atteindre l'étape 3. Considérant les résultats aux t-tests comparant les performances des participants au niveau du hasard, on observe que la performance est nettement au-dessus du hasard pour tout type de structure (N_NV (t (19) = 3.441, p = 0.001), NN_V (t (19) = 3.463, p = 0.001), NV_NV (t (19) = 2.568, p = 0.009), NVN_V (t (19) = 3.034, p = 0.003)). Par ailleurs, il y a un effet principal du match (F (1,19) = 10.808, p = 0.001) ce qui indique que les phrases matching (M = 1.71, ET = 2.41) sont toujours jugés mieux que les phrases mismatching (M = 0.29, ET = 1.23). Deux stratégies concernant l'apprentissage des catégories grammaticales et le nombre sont possibles pour expliquer ces résultats. D'une part, les participants pourraient avoir identifié les deux catégories grammaticales et le nombre, comme supposé dans l'étape 1. D'autre part, il est possible qu'ils aient identifié quatre catégories distinctes. A partir de nos données, il n'est pas possible de départager ces deux alternatives. Cependant, les résultats montrent que les participants sont capables de se concentrer sur la partie pertinente des mots (i.e., les seconde syllabes) et qu'ils ont réussi à identifier les associations correctes entre les syllabes. Cette observation suggère que - quelle que soit la stratégie employée par les participants - elle leur a permis d'identifier les marqueurs morphologiques et de les associer correctement en atteignant l'étape 2. Est-ce que nos participants ont été capables d'atteindre l'étape 3 et apprendre la règle hiérarchique d'accord? Depuis l'analyse de la performance aux phrases matching nous ne pouvons pas distinguer l'étape 2 de l'étape 3 et, par la suite nous ne considérons que la condition mismatching. Nous observons à partir des résultats aux t-test comparant les performances des participants pour les phrases mismatching au niveau du hasard que pour une seule comparaison sur quatre, la performance est nettement supérieure (i.e., NV_NV (t (19) = 1.797, p = 0.044)) ; pour une comparaison, il est marginalement supérieur (i.e., NN_V (t (19) = 1.417, p = 0.086)) et pour deux autres comparaisons (i.e., N_NV et NVN_V (t <1, ns)) la performance n'est pas significativement différente du hasard. Ces résultats suggèrent que les participants n'ont pas atteint l'étape 3 et n'ont pas appris la règle. Fait intéressant, il y a également une interaction entre les variables match et type de connaissances (F (1,19) = 4.133, p = 0.056). Cette interaction reflète le fait que la différence entre matching et mismatching est réduite pour la généralisation lexicale et syntaxique (Matching-Mismatching = 1.16) par rapport à la généralisation lexicale (Matching-Mismatching = 1.69). Ceci est possible parce que la performance pour la généralisation lexicale et syntaxique s'améliore grâce à la structure NV_NV. La structure NV_NV pourrait en fait être traitée comme deux structures NV indépendantes. Comme l'ont montré les analyses statistiques, les participants ont jugé correctement les phrases de structure NV. Ils pourraient avoir appliqué le même principe à la structure NV_NV. Cependant, la connaissance des catégories grammaticales, du nombre et l’association entre les marqueurs morphologiques d’accord est suffisante pour juger ce type de phrase sans l'analyse de la structure en constituants. En outre, le statut particulier de la structure NV_NV - qui pourrait être plus facile que les autres structures - est confirmé par les temps de réaction. Les participants étaient plus rapides à juger les dépendances adjacentes (M = 648 ms, ET = 262) que les dépendances non-adjacentes (M = 724 ms, ET = 341) seulement dans la généralisation lexicale et syntaxique, où les dépendances adjacentes correspondent à des phrases NV_NV.

L'analyse des templates suggère que la condition de matching pourrait avoir empêché les participants d'apprendre la règle correcte. Cette analyse révèle que certains participants semblent
avoir appris une règle linéaire. Comme Endress et collègues (2009) l’ont souligné, la répétition d’un élément joue un rôle important dans l'apprentissage des grammaires artificielles. La présence de ces Primitives Perceptives pourrait avoir influencé le type de régularités que les participants pouvaient apprendre. Dans la présente expérience, les phrases de la condition matching pourraient avoir contribué à l'apprentissage erroné d'une règle linéaire en déviant l'attention des participants vers des propriétés de l’input (c'est à dire, l'identité entre les noms, du-du ou ti-ti) que ne sont pas pertinents à la règle hiérarchique d'accord. Nous observons également que dans la condition matching, les réponses pour la règle hiérarchique d’accord et pour les templates convergent (c'est à dire, quand une phrase est grammaticale pour la règle, elle est grammaticale pour les templates). Il en résulte que, même si les participants ont appliqué une règle incorrecte (i.e., l'un des templates alternatifs) pendant la tâche de jugement de grammaticalité, dans la condition matching ils recevaient toujours des feedbacks positifs. Un tel renforcement positif dans l'utilisation d'une règle incorrecte pourrait avoir empêché les participants d'apprendre la règle correcte.

Expérience 4

Trois changements ont été opérés par rapport à l'Expérience 3. La première modification concerne l'exclusion de la condition matching. En effet dans l'expérience précédente, (la présence de) la condition matching a engendré des difficultés dans l’interprétation des résultats et pourrait avoir empêché les participants d'apprendre la règle d'accord correcte. (En effet, avec la condition matching, les participants pouvaient recevoir des feedbacks positifs en appliquant des règles incorrectes.) Par conséquent, dans la présente expérience, nous avons décidé de n'inclure que la condition mismatching.

Le deuxième changement concerne la variabilité des racines des mots, c'est à dire des éléments qui n'étaient pas impliqués dans la dépendance syntaxique. Dans l'expérience 3, nous avons exploité la variabilité de la première syllabe des mots dans le but de diriger l'attention des participants vers la partie invariante des mots, qui était impliquée dans la règle. Cette idée a été implémentée en utilisant un grand nombre de premières syllabes et seulement quatre possibles syllabes secondes pour construire les mots. Cependant, cette variabilité peut avoir augmenté la charge du traitement lexical, et donc, avoir rendu l'apprentissage de la règle plus difficile. Dans la présente expérience, nous avons décidé de simplifier le traitement lexical afin de favoriser le traitement syntaxique. Pour ce faire, nous avons éliminé les premières syllabes et réduit le lexique à quatre mots monosyllabiques (i.e., du, bo, ti, pa) constitués par les secondes syllabes des mots utilisés dans l'expérience précédente. Par conséquent, le vocabulaire a été constitué par un ensemble beaucoup plus réduit de mots monosyllabiques. Le vocabulaire a été complété par deux autres mots monosyllabiques introduits pour tester la généralisation lexicale.

La troisième modification concerne la segmentation des mots. Nous avons supposé que l'une des raisons pour lesquelles les participants n'ont pas appris la règle hiérarchique d'accord dans l'expérience 3 est qu'ils étaient en train de faire deux tâches en même temps : segmenter le signal acoustique en mots et en extraire les régularités. Dans la présente expérience, nous avons choisi d'ajouter une pause de 25 ms entre les mots. Ce choix vise à faciliter la segmentation et à s'assurer que plus de ressources cognitives seront disponibles et dédiées à l’extraction de la règle. Comme
Peña et ses collègues (2002) l’ont montré dans leurs expériences, les participants sont en mesure d'extraire des structures grammaticales de niveau supérieur uniquement avec un indice de segmentation (i.e., une pause de 25 ms); si cet indice n'est pas disponible, ils restent capable de segmenter les pseudo-mots, mais pas d'extraire la règle qui les caractérise. En conséquence, nous avons ajouté une pause de 25 ms entre les mots et nous avons allongé les pauses entre les constituants à 125 ms afin d'avoir des pauses de longueur différente pour la segmentation entre les mots et entre les constituants.

Dans la présente expérience, nous avons décidé de compléter l'étude de l'apprentissage des règles et la généralisation de la règle avec l'étude de deux autres types de connaissances: la mémorisation et la généralisation syntaxique. Contrairement à l'expérience 3, dans laquelle il était possible de créer de nouveaux mots en introduisant de nouvelles racines tout en gardant les mêmes suffixes, dans l’Expérience 4 le test de généralisation à un nouveau lexique a nécessité la création des nouveaux noms. Deux noms (ge et kE) ont été introduits dans une séance de familiarisation au lexique ainsi que dans une phase d'entraînement impliquant des phrases NV (avec les verbes appris antérieurement). Ainsi, les participants doivent apprendre que le nouveau nom ge appartient à la même catégorie que le nom du et que le nom kE appartient à la même catégorie que le nom ti. Les participants ont ensuite été testés sur des phrases à trois mots. En conséquence, la généralisation lexicale dans l'Expérience 4 diffère de celle de l'Expérience 3, car elle repose sur l'introduction de nouveaux mots qui ne partagent pas de suffixe avec le lexique déjà appris, comme ce fut le cas dans l'expérience 3.

L'expérience consistait en plusieurs sessions: la session de familiarisation au lexique, dans laquelle les participants sont familiarisés avec le lexique, la session d'apprentissage de la règle I, dans laquelle on ils doivent écouter un échantillon de phrases suivant la règle hiérarchique d'accord pour ensuite effectuer une tâche de jugement de grammaticalité. Dans la séance de familiarisation au nouveau lexique, les participants sont familiarisés avec les deux nouveaux noms du lexique, et dans la session d'apprentissage de la règle II, ils ont été invités à généraliser la règle d’accord à des phrases constituées par des deux, trois et quatre mots.

Afin de mieux comprendre les résultats de la présente expérience, il est d'intérêt de porter attention aux différents types de connaissance étudiés dans cette expérience. Pour cette raison, nous avons concentré notre discussion sur les résultats des t-tests comparant les performances des participants à la tâche de jugement de grammaticalité au niveau du hasard. En ce qui concerne la mémorisation, on peut observer que la performance était significativement supérieure au niveau du hasard pour toutes les structures de phrases (NV : t (29) = 8.481, p <.001 ; N_NV : t (29) = 3.642, p <.001 ; NN_V : t (29) = 1.678, p = 0.052). Ces données se réfèrent à la capacité des participants à juger la grammaticalité des phrases sur la base de ce qu’ils ont entendu dans les phases d’entraînement. Les participants pourraient avoir jugé comme correctes les phrases grammaticales uniquement parce qu’ils les reconnaissaient (de la phase d’entraînement) et comme incorrectes les phrases agrammaticales, parce qu'elles étaient nouvelles. Afin d'aller plus loin dans notre compréhension des résultats nous devons regarder les autres types de connaissances testées.
Nous observons que pour la performance à la généralisation syntaxique, les participants ont été meilleurs dans le jugement des phrases caractérisées par la dépendance adjacente (NV_NV (t (29) = 2.769, p = 0.005)) que celle caractérisées par la dépendance non adjacente (NVN_N (t (29) = 1.479, p = 0.075)). Comme déjà mentionné dans la discussion de l'expérience précédente, la structure NV_NV est plus facile que les autres structures. Toutefois, et surtout, la performance était au-dessus du niveau du hasard pour les deux structures. Il est impossible de juger correctement ces phrases sur la base de l'ordre linéaire des éléments, puisque l'ordre grammatical pour une structure (par exemple, ti pa _ bo du) est agrammatical pour l'autre (par exemple, * ti pa du _ bo). Ainsi, ces observations suggèrent que les participants ont appris la règle, ce qui signifie qu'ils doivent avoir intégré l'association entre les deux marqueurs morphologiques et le fait que leur association dépend de la position des mots dans la structure en constituants de la phrase, en fonction de sa prosodie.

Nous considérons maintenant les données de la généralisation lexicale ainsi que de la généralisation lexicale et syntaxique. Nous observons que, après avoir été formés sur les phrases NV avec les deux nouveaux noms, les participants ont réussi effectué la tâche pour les dépendances adjacentes, à la fois sur phrases à trois mots (N_NV (t (29) = 3.052, p = 0.002)) et sur les phrases à quatre mots (NV_NV (t (29) = 2.939, p = 0.003)). Cependant, ils ont échoué dans le jugement des dépendances non adjacents (phrases à trois mots : NN_V : t (29) = -1.606, p = 0.060 ; phrases à quatre mots : NVN_V : t (29) = -0.144, p = 0.443)). Il est possible que parce que les participants ont appris ces mots uniquement avec des phrases NV, ils aient créé une association stricte entre les deux paires de mots (ge - bo et kE - pa), et que par conséquent ils ont jugé les phrases à trois et quatre mots sur la base de la présence (ou absence) de ces paires. Cette hypothèse explique que pour les phrases NN_V la performance des participants se situait nettement au-dessous du niveau du hasard, montrant une tendance systématique à considérer comme correctes les phrases qui contenait l'une des paires, et comme incorrecte celles qui n'en contenaient pas. Cela contraste avec la phase initiale de l'expérience au cours de laquelle les participants ont été formés avec des phrases NV et des phrases à trois mots. La présence de structures à trois mots lors de l'entrainement a limité l'association restreinte par paires entre les noms et les verbes et a souligné l'existence de dépendances non adjacentes, facilitant ainsi leur capacité à étendre la règle aux structures à quatre mots.

En résumé, les résultats de la condition de généralisation syntaxique suggèrent que les participants ont pu apprendre la règle et l'étendre à certaines nouvelles structures. En revanche, les résultats de la généralisation lexicale et de la généralisation lexicale et syntaxique indiquent qu'ils ont échoué à généraliser la règle à des dépendances non-adjacentes avec un nouveau lexique, probablement en raison des conditions d'apprentissage de ce nouveau lexique.

**Expérience 5**

Les résultats de l’expérience 4 suggèrent qu’il est possible d’améliorer l'identification des catégories grammaticales et des marqueurs de nombre, ainsi que l'apprentissage de la relation d’accord entre les marqueurs morphologiques. L’**Expérience 5** a été conçue afin de transmettre ces informations de manière plus saillante. Pour cela nous avons changé deux principales caractéristiques du matériel. La première modification concerne l'utilisation de voyelles finales pour
marquer le nombre (-u pour les mots singuliers et -i pour les mots au pluriel). La seconde modification concerne l'utilisation d'un nombre différent de syllabes pour marquer la distinction entre les catégories grammaticales.

Tandis que dans les expériences 3 et 4 la relation entre les marqueurs morphologiques était basée sur une simple association de cooccurrence, dans l'expérience actuelle, elle est basée sur l'identité : l'accord a été basé sur la répétition d'un même marqueur. Nous avons supposé que la relation d'identité soit plus forte que la relation d'association. En fait, nous pouvons définir l'association comme le lien entre les éléments en mémoire, tandis que l'identité est un cas particulier d'association dans laquelle les éléments sont liés d'une manière perceptuelle et saillante, à savoir par la condition d'être le même. Fait intéressant, il a été démontré que la répétition joue un rôle de facilitation dans l'acquisition de la langue (voir par exemple Endress et al, 2007; Gervain, Macagno, et al, 2008; Mehler et al, 2008; Monaghan & Rowson, 2008 entre autres). Par ailleurs, les premiers mots appris par les enfants sont très souvent caractérisés par redoublement (par exemple, bébé et dodo en Français ou papa (papa) et mamma (maman) en Italien). En outre, plusieurs études avec des adultes ont démontré que la présence de répétitions influence la capacité à généraliser une grammaire à états finis à un nouveau lexique (Endress et al, 2009; Lotz & Kinder, 2006). Si ces répétitions ne sont pas disponibles, la généralisation à un nouveau lexique n'est pas possible (Gomez et al, 2000; Tunney & Altmann, 2001). Par conséquent, les répétitions semblent avoir un statut particulier qui facilite l'extraction et la généralisation des dépendances syntaxiques. En utilisant l'identité plutôt que l'association entre les noms et les verbes pour caractériser l'accord, nous nous attendions à réduire le fardeau cognitif impliqué dans l'apprentissage de l'accord morphologique. Les ressources des participants seront toutes dédiées à la dernière étape d'apprentissage que les apprenants ont à accomplir. (Rappelons que dans cette étape les participants doivent analyser la phrase en constituants, identifier les rôles syntaxiques, et d'apprendre l'accord entre le sujet et le verbe).


En résumé, nous avons émis l'hypothèse que l'ensemble des indices nouveaux introduits dans la présente expérience permettrait d'améliorer la capacité des participants à apprendre et à généraliser la règle hiérarchique d’accord. Le lexique a été constitué de quatre noms au singulier (meku, Rabu, govu, lEpu), quatre noms au pluriel (meki, Rabi, govi, lEpI), deux verbes singulier
(dRu et stu) et de deux verbes au pluriel (dRi et sti). Huit noms supplémentaires (meka, meko, Raba, Rabo, gova, govo, lEpa, lEpo) ont été créés pour le test de vocabulaire.

Une fois encore, l'expérience a consisté en plusieurs séances chacune constituée d'entraînement et test. Cette expérience a été composée de quatre sessions différentes. Dans la familiarisation au lexique les participants ont été familiarisés avec quatre noms du lexique, dans la session d'apprentissage de la règle I, on leur a demandé d'écouter un échantillon de phrases suivant la règle hiérarchique d'accord pour effectuer ensuite une tâche de jugement de grammaticalité. Dans la familiarisation au nouveau lexique les participants ont été familiarisés avec quatre nouveaux noms du lexique, dans la session II de l'apprentissage de la règle, ils ont été invités à généraliser la règle d’accord à des phrases constituées par deux, trois et quatre mots.

Les résultats ont indiqué que, pour ce qui concerne la mémorisation, les participants (N = 20) eu une performance au-dessus de la chance pour l’ensemble des structures étudiées (NV (t (19) = 22.103, p <.001), N_NV (t (19) = 3.752, p <.001), et NN_V (t (19) = 2.680, p = 0.007)). Pour la généralisation lexicale les participants ont eu une performance au-dessus du niveau du hasard pour les phrases N_NV (t (19) = 3.119, p = 0.003) et ils ont également montré une tendance à avoir une performance au-dessus du niveau du hasard pour les phrases NN_V (t (19) = 1.533, p = 0.071). Enfin, pour la généralisation lexicale et syntaxique, nous avons trouvé le même pattern de résultats que dans l'expérience précédente. La performance dans la tâche de jugement de grammaticalité était au-dessus de la chance pour les phrases NV_NV (t (19) = 4.524, p <.001), mais pas pour les phrases NVN_V (t (19) = 0.213, p = 0.416). Encore une fois, lorsque la règle hiérarchique conduit à une dépendance adjacente, les phrases ont été mieux jugées que lorsqu’elle conduit à une dépendance non adjacente. Ce résultat a été mis en évidence par l’effet principal de la variable localité dans l'analyse de la variance (F (1,19) = 11.577, p = 0.003). Considérant ces résultats, nous pouvons raisonnablement penser que les participants ont appris la règle et sa généralisation à un nouveau lexique. Cependant, ils ont échoué quand ils devaient généraliser la règle à la fois à un nouveau lexique et à des nouvelles structures syntaxiques.

Nous considérons maintenant la tendance qui émerge des résultats de l'analyse de templates. L'analyse de templates pour les phrases constituées par quatre mots a suggéré qu'un groupe de six participants a appliqué une règle linéaire pour effectuer la tâche de jugement de grammaticalité (i.e., le troisième et le quatrième mot de la phrase ont le même nombre). Cette observation pourrait expliquer la différence dans les performances entre les structures NV_NV et NVN_V. Les phrases NV_NV pourrait être jugées en appliquant cette règle linéaire. Cependant, ce n'était pas le cas pour les phrases NVN_VL'analyse de templates a également montré qu'il y avait un groupe de sept participants qui ont réussi à apprendre la règle et à la généraliser à un nouveau lexique (voir l'analyse de templates pour les phrases à trois mots), et un autre petit groupe de cinq participants qui a généralisée la règle à un nouveau lexique et à des nouvelles structures syntaxiques (voir l'analyse de templates pour les phrases à quatre mots).

Les performances des participants semblent être très hétérogènes. Cependant, nous pouvons raisonnablement proposer de les classer en fonction de différents profils. 1) Un premier profil se réfère aux participants qui n’ont pas appris la règle ou tout autre template. Les participants
appartenant à ce groupe n'ont pas atteint le niveau du hasard pour la règle ou l'un des templates proposés. 2) Le deuxième profil se réfère aux participants qui ont appris une règle différente de la règle hiérarchique d'accord. Ces participants ont eu une performance au-dessus du hasard pour l'un des templates alternatifs. 3) Le troisième profil se réfère aux participants qui ont appris la règle, mais qui ont réussi à la généraliser seulement à un nouveau lexique. Ces participants sont ceux qui ont une performance au-dessus du hasard avec la règle pour les phrases de trois mots, mais pas pour les phrases de quatre mots. 4) Le quatrième profil se réfère aux participants qui ont eu une performance au-dessus de la chance pour la généralisation lexicale et syntaxique (phrases de quatre mots). Nous pouvons raisonnablement penser que ces participants, bien que plus lent que d'autres, ont appris la règle. 5) Enfin, le dernier profil se réfère à des participants qui ont appris et généralisé la règle à un nouveau lexique et à de nouvelles structures syntaxiques. Ces participants ont eu une performance au-dessus du hasard pour la règle, tant pour les phrases de trois mots et les phrases de quatre mots.

L'Expérience 5 apporte de meilleures conditions d'apprentissage de la règle hiérarchique d'accord en comparaison avec l'Expérience 4, en particulier pour ce qui concerne les phrases NN_V. Il est ainsi vraisemblable que les nouveaux indices introduits (indice phonologique différent pour marquer la catégorie grammaticale et identité entre les voyelles finales pour marquer le nombre) ont permis d'améliorer l'apprentissage de la règle. D'autres changements mineurs avaient également été introduits dans la présente expérience (par exemple, des pauses plus longues aux frontières des constituants, le contour mélodique commun pour les structures NV intégré dans d'autres phrases). En conséquence, nous ne pouvons pas déterminer avec certitude quelle(s) variable(s) a joué le rôle principal dans cette amélioration. Cependant, nous pouvons raisonnablement penser que les nouveaux indices, et surtout le fait que l'accord morphologique a été déterminé par la relation d'identité entre les voyelles finales, a eu un impact fondamental sur la performance des participants. La saillance de la relation d'identité par rapport à la simple association utilisée dans l'expérience précédente pourrait avoir joué le plus grand rôle dans l'apprentissage.

LE ROLE DE LA SEMANTIQUE

Sur la base des résultats des expériences précédentes, nous avons conclu que, alors que l'apprentissage et la généralisation d'une règle hiérarchique d'accord sujet-verbe dans le contexte d'une langue artificielle ne soit pas une tâche facile, cette tâche semble possible quand l'ensemble des indices nécessaire sont fournis de façon appropriée. Puisque nous souhaitons étudier la manière dont on peut apprendre de telles règles avec un degré de maîtrise plus élevé, nous avons décidé d'ajouter un indice. Ainsi dans l'Expérience 6, nous avons décidé d'investiguer le rôle de la sémantique et étudierons la saillance de cet indice dans l'apprentissage de la règle hiérarchique d'accord dans un paradigme d'apprentissage des langues artificielles. (Pour rappel, dans les expériences précédentes, la pertinence des différents indices phonologiques pour marquer la catégorie des mots et celle des indices prosodiques pour marquer la structure constitutive de la phrase ont été explorés.) À notre connaissance quelques études seulement (Moeser & Bregman, 1972; Morgan & Newport, 1981; Morgan et al, 1987; Mori & Moeser, 1983) ont analysé le rôle de la sémantique dans l'apprentissage de règles syntaxiques dans un paradigme d'apprentissage d'une
langue artificielle. Dans d'autres études, les chercheurs, principalement intéressés définir des corrélats corticaux de l'apprentissage des langues et du traitement syntaxique, ont inclus la sémantique dans leurs langues artificielles. Ils ont supposé que la sémantique faciliterait l'apprentissage (voir par exemple Amato & MacDonald, 2010; Friederici, Rüschmeyer, Hahne, & Fiebach, 2003; Friederici, Steinhauser, & Pfeifer, 2002; Opitz & Friederici, 2004), bien qu'ils n'aient pas directement étudié son rôle.

Deux études (Moeser & Bregman, 1972; Morgan & Newport, 1981) sont particulièrement pertinentes dans le cadre de notre expérience. Moeser et Bregman (1972) ont testé la capacité des adultes à apprendre une grammaire syntagmatique dans différentes conditions variant selon les références sémantiques. Ils ont utilisé des caractéristiques visuelles (par exemple, formes, lignes, orientation) pour donner une référence à des mots dépourvus de sens et pour tester l'apprentissage de la grammaire artificielle. Leurs résultats ont montré qu'une grammaire complexe a été apprise seulement quand il y avait une correspondance entre la syntaxe et la sémantique, c'est-à-dire quand les dépendances syntaxiques entre les mots ont été marquées par la combinaison de caractéristiques visuelles. La précision de réponse a été d'environ 79% quand il y avait une correspondance entre syntaxe et sémantique, alors qu'elle avoisinait les 47% quand il n'y avait qu'une correspondance arbitraire entre les mots et les caractéristiques visuelles. Les auteurs ont conclu que l'apprentissage d'une grammaire complexe n'était possible que lorsque les dépendances syntaxiques sont représentées sémantiquement.

D'autre part, Morgan et Newport (1981) font l'hypothèse que ce n'était pas la sémantique qui a permis aux participants de l'étude de Moeser et Bregman (1972) d'apprendre la grammaire, mais la présence d'indices marquant les frontières syntaxiques. Dans leur étude, le groupe des participants bénéficiant de l'indice de regroupement en constituant mais pas d'indices sémantiques a appris avec succès la grammaire (précision d'environ 84%) et leur performance ne différait pas significativement de celle du groupe des participants bénéficiant aussi d'indices sémantiques en termes de précision (90%). Ces résultats soutiennent l'hypothèse selon laquelle lorsque les indices marquant l'organisation des phrases en constituants sont disponibles dans l’input, les participants apprennent avec succès une grammaire complexe. Toutefois, une différence a été observée dans les tendances d'apprentissage des deux groupes de participants, suggérant des stratégies d'apprentissage différentes. Ces résultats n'excluent donc pas totalement la possibilité d'une facilitation de la sémantique. Nous avons observé que certaines données de la littérature et les résultats de l'Expérience 5 vont à l'encontre de la nécessité de la sémantique pour apprendre la grammaire. Cependant, comme nous l'avons mentionné ci-dessus, il est possible que la sémantique joue un rôle facilitateur dans son acquisition. L’Expérience 6 visait à tester cette hypothèse.

Expérience 6

Le matériel et la procédure et le nombre de participants (N = 20) de cette expérience étaient identiques à ceux de l'expérience 5. L’unique différence concerne l'association d'une vidéo à la phrase lors des phases d’entraînement. Alors que dans l'expérience 5, la vidéo était toujours le même (un objet en mouvement avec une trajectoire aléatoire), dans l'expérience 6 chaque phrase a été associée à une vidéo différente. Chaque nom a été associé à une forme géométrique colorée (par
exemple, un carré, un cercle, un ovale, une étoile) tandis que les verbes ont été associés à des actions (par exemple, déplacer ou tourner). Les noms singuliers étaient représentés par une forme géométrique (par exemple, un carré, un cercle, etc.) tandis que les noms pluriels ont été représentés par un groupe de trois formes géométriques identiques (par exemple, trois carrés, trois cercles, etc.). Nous avons observé que les participants ont eu des scores au-dessus de la chance pour tous les types de connaissance (i.e., la mémorisation, la généralisation lexicale, la généralisation lexicale et syntaxique) et pour toutes les structures de phrases étudiées (Mémorisation: NV (t (19) = 16.692, p <.001), NN_V (t (19) = 3.164, p = 0.002), N_NV (t (19) = 2.046, p = 0.027); Généralisation lexicale: NN_V (t (19) = 3.177, p = 0.002), N_NV (t (19) = 4.020, p <0.001); Généralisation lexicales et syntaxiques: NV_NV (t (19) = 3.821, p <.001), NVN_V (t (19) = 1.984, p = 0.031)). Cela indique que les participants ont pu apprendre la règle et la généraliser à la fois à un nouveau lexique et à de nouvelles structures syntaxiques.

Les résultats de l'analyse de variance ont montré que l'avantage pour les dépendances adjacentes (M = 2.1, ET = 2.46) par rapport aux dépendances non adjacentes (M = 0.65, ET = 1.47) était présent pour la généralisation lexicale et syntaxique, mais absent pour la mémorisation et la généralisation lexicale. La seule différence observée entre dépendances adjacentes et non adjacentes pourrait être due au fait que les participants éprouvent plus de difficultés pour la structure de NVN_V soit que pour la structure NV_NV. Dans les expériences précédentes, nous avons en effet constaté que les participants ont toujours eu de meilleurs performances avec la structure NV_NV en comparaison avec les autres structures, possiblement parce que la structure NV_NV est traitée comme deux structures NV indépendantes.

Comme dans l'expérience 5, nous pouvons identifier quatre profils de performance différents dans cette expérience. 1) Le premier profil renvoie à des participants qui n'ont pas appris la règle ou tout autre template alternatif. 2) Le second se réfère aux participants qui ont appris un autre template. L'analyse des phrases de quatre mots a révélé qu'il y a trois participants qui ont employé le template 3<sup>3</sup>-4<sup>-4</sup>. 3) Les participants appartenant au troisième profil ont appris la règle et la généralisation à un nouveau lexique. 4) Les participants appartenant au quatrième profil ont appris la règle et la généralisation à la fois à un nouveau lexique et à des nouvelles structures syntaxiques.

La comparaison entre l'Expérience 5 et l'Expérience 6 visait à éclaircir la contribution de la sémantique dans l'apprentissage d'une règle d'accord hiérarchique. L'hypothèse que nous avons testée était de savoir si la sémantique améliore considérablement l'apprentissage des langues artificielles. Les résultats de l'analyse de variance sur les valeurs d-prime ont souligné qu'il n'existe pas de différence significative entre les performances des participants à l'Expérience 5 et à l'Expérience 6, i.e., l'effet principal de la variable Expérience n'était pas significative : F < 1, ns. Cette observation suggère que la sémantique ne facilite pas l'apprentissage de la règle d'accord hiérarchique dans notre langue artificielle. Toutefois, si on regarde les résultats aux t-tests comparant les scores des participants au niveau du hasard pour chaque type de connaissance et pour chaque structure de la phrase, une différence apparaît. Dans l'Expérience 5 seule la performance pour la structure NV_NV dans la généralisation lexicale et syntaxique était significativement supérieure au hasard alors que dans l'Expérience 6, la performance pour les deux structures de
phrase à quatre mots était significativement supérieure au hasard. Ces résultats indiquent que l'apprentissage est bon dans les deux expériences, mais ils suggèrent également que la présence de la sémantique permette la généralisation de la règle à un nouveau lexique et à de nouvelles structures syntaxiques plus complexes. Nous pouvons supposer que la sémantique joue un rôle seulement quand un traitement de niveau supérieur est requis plutôt qu'une simple mémorisation. Fait intéressant, lorsque nous avons considéré dans l'analyse de la variance seulement les conditions de généralisation lexicale et de généralisation lexicale et syntaxique, la valeur p de l'effet principal de la variable expérience n'était pas très élevé (p = 0.149). Il est raisonnable de supposer que l'absence de différence statistiquement significative entre les deux expériences est due à la forte variabilité des performances des groupes. Par conséquent, on peut penser que l'augmentation de la puissance statistique (i.e., le nombre de participants) conduirait à une différence significative entre les expériences avec un avantage pour celle comportant la sémantique.

Comme la généralisation lexicale et syntaxique a été testée avec les phrases à quatre mots, on peut obtenir des renseignements supplémentaires au sujet des différences entre l'Expérience 5 et l'Expérience 6 en regardant l'analyse des templates. Dans l'Expérience 6, neuf participants ont effectué la tâche au-dessus de la chance avec la règle, tandis que dans l'Expérience 5 seulement cinq participants ont effectué la tâche au-dessus de la chance avec la règle. Cette observation suggère qu'il y a eu une augmentation dans le nombre de participants qui ont appris avec succès la règle pour ce type de connaissances. Par ailleurs, dans l'Expérience 6 seulement trois participants ont effectué la tâche au-dessus de la chance avec un autre template, tandis que dans l'Expérience 5, ils étaient six. Cette observation suggère que moins de participants pourraient avoir appliqué une autre règle. Par conséquent, nous pouvons raisonnablement conclure qu'il y a une influence de la sémantique quand l'apprentissage implique un type de traitement supérieur, comme celui permettant la généralisation de la règle à la fois à un nouveau lexique et à des nouvelles structures syntaxiques.

**DISCUSSION GENERALE**

Le but de cette thèse était double. Tout d'abord, nous voulions tester la capacité des adultes à apprendre les règles d'accord dans une langue artificielle, deuxième, nous voulions étudier les facteurs influençant cet apprentissage. Nous nous sommes centrés sur l'apprentissage de deux types de dépendances syntaxiques: dépendances linéaire et hiérarchique.

Les résultats des expériences 1 et 2 montrent que les participants n'ont pas appris la règle linéaire d'accord dans le domaine de la musique ou du langage. En particulier, dans le domaine de la musique les participants n'ont été pas capables d'apprendre la règle proposée ou une règle alternative. Dans le domaine linguistique, les participants pourraient être classés en deux groupes selon leurs performances. Les participants appartenant au premier groupe n'ont pas appris la règle d'accord ou une autre règle, tandis que les participants appartenant au second groupe ont appris l'une des règles alternatives (c'est à dire, le premier, le deuxième et le troisième élément de la phrase ont le même nombre). Nous supposons que la saillance perceptive d'une catégorie particulière d'items (i.e., phrases matching grammaticales) a induit en erreur les apprenants.
Dans les expériences 1 et 2, nous avons employé différents types d'unités pour la construction de la langue artificielle (mots-ton ou pseudo-mots), afin de déterminer si ce facteur influe sur l'apprentissage. Cependant, parce qu'aucune expérience n’a révélé d'apprentissage, il n’est pas possible de répondre à cette question.

Les résultats obtenus dans les Expériences 3 à 6 sont très hétérogènes; les participants n'ont pas les mêmes performances dans toutes les étapes d'apprentissage ni pour tous les types de connaissance. En particulier, nous n'avons pas observé d'apprentissage de la règle dans l'Expérience 3, dans lequel le lexique a été constitué par des pseudo-mots bisyllabiques et dans laquelle les catégories et le nombre ont été marqués par des caractéristiques phonétiques des consonnes et des voyelles. Dans cette expérience, les participants ont pu apprendre la distinction entre les catégories grammaticales et l'association entre les marqueurs d’accord. Dans l'Expérience 4, dans laquelle le lexique a été constitué par des pseudo-mots monosyllabiques et dans laquelle les catégories et le nombre ont été marqués avec des caractéristiques des différentes consonnes, nous avons observé l'apprentissage de la règle, mais aucune généralisation à nouveau lexique ni à des nouvelles structures syntaxiques. Dans les Expériences 5 et 6, dans lesquelles les catégories ont été marquées par un nombre différent de syllabes et le nombre par les voyelles finales, les participants ont pu apprendre et généraliser la règle avec un bon niveau de maîtrise. En regardant plus attentivement les performances individuelles, nous avons observé l'existence de différents profils de performance. À un extrême, les participants n'ont pas réussi à apprendre la règle et à l'autre extrême, les participants ont réussi à apprendre la règle et sa généralisation. Par ailleurs, certains participants ont appris l'une des règles alternative, et d'autres participants ont bien réussi sur un type de connaissance uniquement (par exemple, la généralisation lexicale mais pas la généralisation lexicale et syntaxique).

De façon générale, nous pouvons conclure (de ces résultats) que les adultes sont capables d'apprendre des dépendances syntaxiques hiérarchiques dans des langues artificielles. Dans les Expériences 5 et 6, certains participants ont acquis la règle d'accord hiérarchique et ils ont été capables de la généraliser à des nouvelles phrases. Ce résultats nous conduisent ainsi à répondre «oui» à notre première question de recherche. Par conséquent, la question principale n'est plus de savoir si les adultes sont capables d'apprendre des dépendances syntaxiques hiérarchiques, mais quels sont les facteurs qui influencent cet apprentissage, ce qui nous introduit à la deuxième conclusion.

La seconde conclusion que nous pouvons tirer des présents résultats est que certains indices phonologiques sont plus efficaces que d'autres pour marquer des informations telles que les catégories grammaticales et le nombre. En effet, les participants ont obtenues de meilleures performances dans l'Expérience 5, où les catégories grammaticales étaient marquées par le nombre de syllabes, que dans l'Expérience 4, où les catégories grammaticales étaient marquées par le point d'articulation de certains phonèmes dans les mots. Nous pouvons supposer que l'utilisation du nombre de syllabes pour catégoriser les mots est plus efficace que le point d'articulation des phonèmes. Cela n'est pas surprenant, puisque le nombre des syllabe est l'une des propriétés phonologiques qui différencie les catégories grammaticales (Kelly, 1992, 1996) dans le langage naturel. Par ailleurs, la performance des participants dans la tâche d'apprentissage a nettement
augmenté lorsque l'accord morphologique était marqué par l'identité de voyelles finales du sujet et du verbe, plutôt que lorsqu'il était simplement marqué par la cooccurrence de certaines syllabes. Cette observation est pertinente pour les modèles théoriques de l'acquisition du langage naturel. En effet, elle peut être interprétée comme une preuve du statut particulier de la relation d'identité. Ce statut est reconnu notamment par l'approche des primitives perceptuelles pour l'acquisition du langage.

En ce qui concerne les indices prosodiques nous sommes arrivés à deux conclusions. Tout d'abord, l'amélioration des performances des participants dans l'Expérience 4 par rapport à l'Expérience 3 montre qu’ils ont bénéficié de la présence des pauses entre les mots. Ces pauses ont facilité la segmentation du signal et réduit la charge cognitive qui pouvait résulter du fait que dans l'Expérience 3 les participants devaient effectuer deux tâches en même temps (la segmentation et l'apprentissage des règles). Deuxièmement, en utilisant des pauses plus longues entre les constituants, nous avons eu une meilleure mise en œuvre de la règle hiérarchique. Nos résultats sont en ligne avec les interprétations selon lesquelles l'extraction de certains types de dépendances syntaxiques dépend de la présence d'indices de segmentation dans l’input (Morgan & Newport, 1981; Morgan et al, 1987).

Le rôle de la sémantique dans le traitement du langage naturel a été largement étudié, mais seules quelques études l’ont directement exploré dans les langages artificiels (Moeser & Bregman, 1972; Morgan & Newport, 1981). Nous avons contribué à la récolte de preuves dans ce dernier domaine en comparant les Expériences 5 et 6. Dans l'Expérience 6 nous avons étudié l'influence des indices sémantiques en créant une correspondance entre les noms et des formes géométriques, et entre les verbes et des actions. La comparaison des résultats de cette expérience avec les résultats de l'Expérience 5 (pour laquelle la seule différence porte sur l'absence d'indices sémantiques) suggère que: 1) les indices sémantiques ne sont pas nécessaires contrairement à ce qu’avaient proposé par Moeser et Bregman (1972), puisque l'apprentissage était déjà observé dans l'Expérience 5 et qu’il n'y a pas de différence significative entre les performances des participants aux deux expériences; 2) Il est néanmoins possible que les indices sémantiques aient une influence sur l'apprentissage, car il y avait, dans l'Expérience 6, plus de participants qui apprenaient la règle et moins qui apprenaient des règles alternatives que dans l'Expérience 5.

Cette thèse représente une première tentative pour étudier une langue artificielle plus "naturelle" que celles généralement étudiés dans les paradigmes ALL. Plus de d'études sont nécessaires pour comprendre de façon plus approfondie la capacité d’apprendre des dépendances syntaxiques hiérarchiques et pour caractériser la représentation et le traitement de ces structures.