SCUOLA DI DOTTORATO
UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA

Department of Psychology

PhD program in Psychology, Linguistics and Cognitive Neurosciences
Cycle XXX
Curriculum mind, brain and behaviour

PROCESSING AND LEARNING OF SEQUENTIAL PATTERNS IN DEAF AND HEARING INDIVIDUALS: DIFFERENCES AND SIMILARITIES.

Surname: **Giustolisi** Name: **Beatrice**
Registration number: **798822**

Supervisor: **Prof. Carlo Cecchetto**

Co-supervisor: **Prof. Maria Teresa Guasti**

ACADEMIC YEAR 2016/2017

In loving memory of Francesco, mio zio.

Abstract

The Auditory Scaffolding Hypothesis (ASH, Conway et al., 2009) suggests that sound experience provides a scaffolding for the development of sequencing and timing behavior. As a result, a period of auditory deprivation from birth should imply sequencing and timing behavior impairments. The socio-clinical consequences of this hypothesis are not trivial. In fact, considering that sequencing and timing behavior are among the building blocks of many cognitive functions (see Lashley, 1951), if the ASH is correct swift action should be taken to ensure each deaf newborn to have his/her sense of hearing restored with adequate technologies. In view of the above, I considered crucial to test predictions that follow from the ASH to verify its validity.

My dissertation is composed by five chapters: an introductory chapter (1), three experimental chapters (2, 3 and 4) and a final summary (5). Participants of the studies reported in chapters 2 and 3 were Deaf adult signers, whereas in chapter 4 deaf children with cochlear implant (CI).

In **chapter 1**, I introduce the ASH in great detail, explaining its theoretical basis. Then, I highlight some criticisms. The chapter contains also a digression on the nature of sequencing and timing behavior following the taxonomy described in Dehaene et al. (2015). I explain that assessing sequencing and timing behavior should take as a reference point the Dehaene taxonomy, and that my agenda consists indeed of the assessment of the ASH at different levels of this taxonomy.

Chapter 2 reports a set of four experiments assessing the ability of Deaf and hearing adults to acquire artificial grammars presented in the visual modality. Even if the strategy adopted by the two groups of participants might differ, Deaf as well as hearing adults were able to acquire artificial grammars at different levels of the Chomsky hierarchy, from finite state grammars to context-sensitive grammars.

Chapter 3 presents a study on visual statistical learning with sequentially presented stimuli in Deaf and hearing adults. Moreover, this study investigated the relationship between visual statistical learning skills and literacy. Visual statistical learning scores positively correlated with reading comprehension scores in the hearing group. The correlation was also positive in the Deaf group, but it was not significant. Even if the Deaf group had lower reading skills than the hearing group, the two groups exhibited comparable visual statistical learning abilities.

Chapter 4 contains a study on timing abilities in children with CI. Children's ability to construct the abstract representation of regular isochronous stimuli was assessed using a warning-imperative task. Participants' results were extremely variable. Nevertheless, the vast majority of children performed within standard limits.

The last chapter, **chapter 5**, is a summary of the arguments and evidence against the ASH that I presented in the previous chapters.

To sum up, in this dissertation I argue that the ASH cannot be accepted as valid hypothesis and that sequencing and timing behavior can fully develop also in people with delayed/no access to auditory experience. Therefore, input other than sound should sustain these skills, which are among the building blocks of many high-level cognitive functions.

Acknowledgments

In the last three years, being a PhD student made me travel around Italy, Europe and the US. I spent time chatting about linguistics in the magnificent squares of Rome, I learnt about the neurobiology of language in the Dutch city of Nijmegen and in Long Island, New York. I studied statistics in the little and pretty Bertinoro, in the center of Italy, and I learned about Sign Languages by the canal of Venice and in Barcelona. I collected data in Milano, Torino, Monza, Piacenza, Modena, Vienna, and San Diego. And it is in San Diego that I wrote more than a half of this dissertation, under a beautiful blue sky. Being in so many different places provided me with a big amount of inputs, which made the learning experience that characterize my PhD so special.

Of course, it is not just about the places. It is more about the people that I met and that I would like to thank for their contribution in making me growing as a researcher, and as a person.

First, I would like to thank my supervisor, Professor Carlo Cecchetto for guiding me through the PhD, for all the opportunities he gave me, and all the opportunities he is giving me. Thank you Carlo, I consider myself very lucky to arrive in Milano before your departure for Paris.

Then, I would like to thank my co-supervisor, Professor Maria Teresa Guasti, for sharing with me so many ideas and for believing in me. Thank you, your support meant a lot to me.

A heartfelt thanks goes to all BIL group members, to le Francesche

and Fabrizio for the work we did together, to Mirta for discussing with me about statistics, to Elena for sharing with me her experience, to Lara for teaching me about Sign Languages and to all the other past and present members. I would also thank Professor Natale Stucchi for getting me involved in many exciting projects.

A big thank goes to my supervisor while I was in San Diego, Professor Karen Emmorey. Thank you and thank to all LLCN members and especially to Cindy, who helped me with data collection.

I also would like to thank Dr. Gesche Westphal-Fitch and Professor Tecumseh Fitch for hosting me in Vienna and teach me about artificial grammar learning.

Then, I thank my dissertation reviewers, Prof. Naama Friedman and Dr. Chloe Marshall for their precious comments on the first version of my dissertation.

A very big thank goes also to my colleagues of the XXX PhD cycle, for sharing with me this three incredible years, a countless number of whatzapp messages, and e-mails, and cups of coffee, doubts, suggestions, assistance. Thank you, my friends! A special thank to Alessandra, my favorite co-author, and to Robi B and Marco for being such a support during our stay in the Pacific Time Zone.

Being a PhD student was a priceless experience, even though sometimes it was hard. I cannot count the times I was unsure about my chance to get to the end. I was lucky enough to be surrounded by many people who gave me strength, especially at home.

I would like to thank my parents, my biggest supporters, for the unconditional love they give me.

My last, and deepest thank goes to Paolo, my life partner, for encouraging me when being a PhD student was hard and for reminding me that unplugging is vital.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

This first introductory chapter presents a hypothesis about possible effects of auditory deprivation on cognitive development, and specifically the consequences on the development of temporal and sequential skills. I discuss the first critical thought I had on this hypothesis, and the relevance of those skills for human behavior. Then, I propose a framework to analyze and assess the before mentioned hypothesis, which is the main goal of this dissertation. I present some work arguing against this hypothesis and I outline the experimental studies contained in this dissertation.

1.1 The Auditory Scaffolding Hypothesis

Sensory systems process specific forms of energy to ultimately guide and modify behavior (Purves et al., 2008). Somehow my dissertation has to do with one of the most important sensory system for humans, and exactly with hearing and the perception of auditory stimuli. To be more precise, what follows deals with an impaired sense of hearing, and a subsequently highly reduced or absent (and in some cases restored) perception of auditory stimuli. To be even more precise, the wider question behind the studies collected in this dissertation is what is the relation between hearing loss and sequencing behavior.

My interest in this specific topic comes from a precise hypothesis stating that humans need hearing to develop sequencing skills. The so called Auditory Scaffolding Hypothesis (hereinafter ASH) was first proposed by Christopher M. Conway, David B. Pisoni and William Kronenberger in a 2009 paper entitled *"The importance of Sound for Cognitive Sequencing Abilities"*. According to the ASH:

> [...]Sound is inherently a temporal and sequential signal. Experience with sound therefore [...] provide a kind of "scaffolding" for [...] the development of general cognitive abilities related to representing temporal and sequential patterns (p. 275)

From a physics standpoint, temporal and sequential dimensions can indeed be considered the foundations of sound stimuli (see Figure 1.1). In fact, four elements are needed to describe a sound wave: the waveform, the amplitude, the phase and the frequency, which is expressed in cycles per second (or Hertz, Hz).

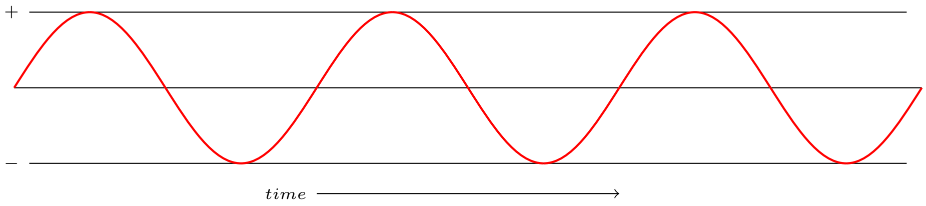


Figure 1.1: A schematic representation of sound waves.

From the psychophysics and behavioral perspective, the two main pieces of evidence in favor of the ASH were i) that hearing people perform better in rhythmic, working memory and sequence learning tasks when the stimuli are auditory instead of non-auditory, and ii) that hearing-impaired populations seem to be impaired in sequential tasks. I will examine these two pieces of evidence in the next paragraphs.

A well replicated effect in psychophysics is that hearing adults synchronize better to auditory than to visual stimuli. Specifically, when people are asked to tap their index finger in synchrony with isochronous auditory or visual sequences (i.e., sequences of tones or visual flashes), the performance is worse for the visual modality than for the auditory modality (e.g. Repp and Penel, 2002). Moreover, auditory distractors when tapping on a visual sequence interfere with the performance, whereas visual distractors when tapping on an auditory sequence have almost no effect (Repp and Penel, 2004). As for working memory, short-term verbal memory is known to be higher for verbal stimuli (i.e. spoken words) than for visual stimuli (i.e. written words) (Penney, 1989), and span is larger for speech than for signs also in balanced bimodal bilinguals (Boutla et al., 2004; Bavelier et al., 2008). Finally, hearing people learn sequential patterns better when they are made of auditory stimuli instead of tactile or visual stimuli (Conway and Christiansen, 2005). To sum up, better performance with auditory stimuli in synchronization, working memory and sequential learning tasks might actually suggest that temporal and sequential patterns are better suited for the auditory modality.

The second piece of evidence in favor of the ASH, i.e. the observation of sequence learning impairments in deaf children, was mainly grounded on two studies. In the first study, a group of deaf children with cochlear implant (CI), all born in hearing families, and a control group of hearing children matched in age took part in an implicit visual sequence learning task of visual non-linguistic stimuli. Participants were asked to memorize and reproduce a sequence of colored squares presented one by one on a computer screen. Without informing the participants, the task was divided into two phases, a familiarization (or learning) phase followed by a testing phase (cf. chapter 2 and chapter 3 for more information about this paradigm). During the learning phase, the sequence of colors followed some specific constraints, and specifically transition probabilities between colors were fixed. During the testing phase, half of the sequences followed the constraints of the learning phase (familiar stimuli), and half of the sequences did not (unfamiliar stimuli). Implicit sequence learning was assessed comparing accuracy in reproducing familiar stimuli with accuracy in reproducing unfamiliar stimuli.In fact, the difference in accuracy between the two types of testing stimuli should indicate if implicit learning has occurred. The control group showed a learning effect, whereas the deaf children did not. Performance on the implicit learning task correlated with language skills (Conway et al., 2011b). Moreover, the same group of children with CI was tested on a motor sequencing task, the NEPSY "finger tapping" test (Korkman et al., 1998), which requires to tap the index finger against the thumb separately for each hand, and all the fingers sequentially against the thumb. Scores on this tasks are derived from the time required for task completion. Children with CI showed impaired performance compared to hearing peers (Conway et al., 2011a).

These results seem therefore to confirm the ASH. Lack of hearing experience from birth appears to interfere with a set of cognitive tasks that involve learning and production of sequences. Nevertheless, some criticism might be raised: I will discuss some of these criticisms in the next section.

Before continuing, I would like to summarize again the ASH, reproducing its schematic representation, provided in Figure 2 of Conway et al. (2009) (Figure 1.2 here). I will get back on this schema in chapter 5.

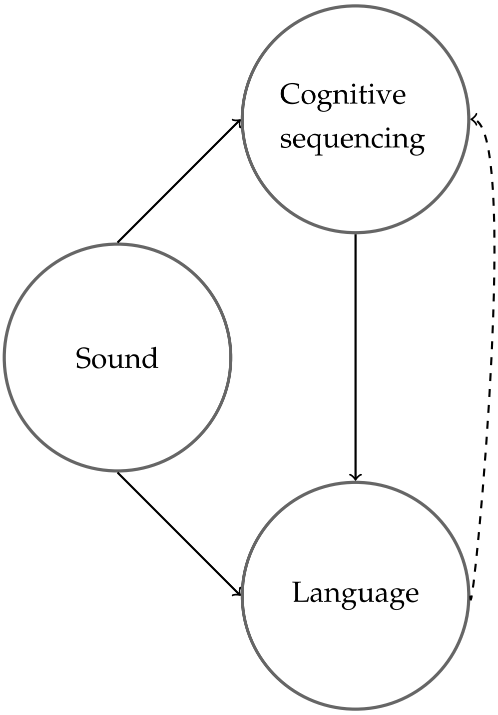


Figure 1.2: Schematic representation of the ASH. Redrawn from Conway et al. (2009).

Sound affects Cognitive Sequencing and Language, Cognitive Sequencing affects Language. The curved dotted line from Language to Cognitive Sequencing represent, in Conway's words "an additional but as-yet unspecified influence: spoken language skills affecting the development of general cognitive sequencing abilities".

1.2 Criticisms

This paragraph considers the results reported in Conway et al. (2011b,a) as a proof for the ASH. Before trying to find the source of the impairment in learning/producing sequences in the lack of auditory experience (which, by the way, had been to some extent restored trough CIs in the above mentioned studies), one should consider the full status of the participants. Specifically, the deaf participants of the two Conway's studies (Conway et al., 2011a,b) were born in hearing families. In addition to auditory deprivation, these children have probably underwent a period of language deprivation, plus their learning process of a spoken language was mediated by speech therapy, at least during the first phases. Moreover, the sociocultural environment in which those children were raised is problematic. As pointed out by Hauser et al. (2008), hearing parents might initially consider as a tragedy the birth of a deaf child, and even when this is not the case, they might require some time to develop the skills that are needed to raise a deaf child. Therefore, impaired performance in implicit sequential learning might be the result of a more complex developmental situation, in which auditory deprivation interacts with language exposure delays and a delicate family environment. The ASH should indeed be tested considering also deaf children of Deaf parents [1], auditory deprived from birth, but not language deprived since exposed to a natural language (a signed language) from birth. As Marschark and Hauser wrote: "The issue of language clearly is one that is woven throughout our understanding of cognition, learning, and the development of deaf..." (2008). And not to be neglected, Deaf parents are presumably prepared for the chance to have a deaf child, which should result in a less critical domestic situation. Moreover, to confirm the validity, the ASH should be tested with adult participants, to see if deaf people eventually show the same sequence skills as hearing adults.

As for the modality-specific constraints in several sequences/timing-related cognitive tasks in hearing populations, recent work suggested that also ultra-replicated experimental findings like the auditory advantage in motor synchronization are more experience- and stimulus depen-

---

[1] I will use the term *Deaf*, with capital D, to refer to the deaf people that recognize themselves as a member of the Deaf community, characterized by a specific language (a sign language) and cultural identity (Woodward, 1972).

dent than has ever been thought. Namely, it has been proved that Deaf individuals synchronize to visual flashes better than hearing people do, and that hearing people are able to synchronize to moving visual stimuli (i.e. to a bouncing ball) as accurately as to auditory stimuli (Iversen et al., 2015). I will go back to this in great detail in chapter 4.

To sum up, although the ASH was supported by several different pieces of evidence, from my point of view none of these is entirely convincing. Before continuing, it is necessary to make a digression concerning the nature of cognitive sequencing abilities.

1.3 About sequences

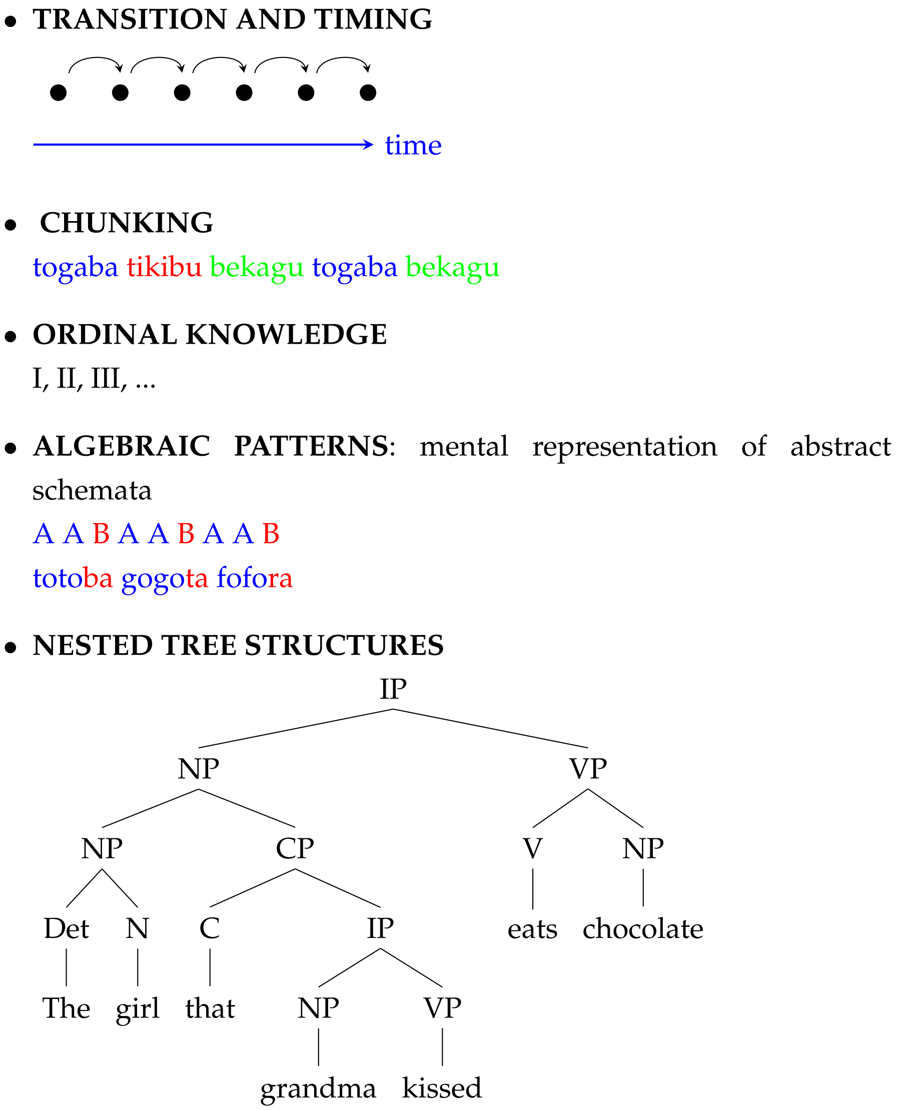
Why are sequences important? The answer is straightforward: Because sequences are among the building blocks of many cognitive functions, from working memory to action planning, and language processing as well."Not only speech" said Lashley (1951) "but all skilled acts seem to involve the same problem of serial ordering, down to the temporal coordination of muscular contractions in such a movement as reaching and grasping". Let's suppose I am thirsty, and I decide to drink some water from the bottle standing next to my laptop. I should first open the bottle, and then bring the bottle to my mouth and tilt it. Open the bottle while it is tilted towards me would probably result in me getting wet, before I could eventually drink the spared water.

This is one point that I would like to emphasize: the importance of sequences for human development makes the socio-clinical consequences of the ASH not trivial. As stressed also by Hall et al. (2017), if the ASH is correct, and humans need sound input to develop sequencing abilities, swift action should be taken to ensure each deaf newborn to have his/her sense of hearing restored with adequate technologies (e.g. CIs), and this should be true regardless of how parents/caregivers have chosen to communicate with their children. On the contrary, if the ASH is not correct, and inputs other than sound (a sign language, for example) contribute to the development of sequencing abilities, the importance of restoring hearing may be contingent on the exposure of these other input sources, which, in the case of language, means either signed or spoken language (or both).

1.3.1 A taxonomy for sequences representation

Sequences can be coded at least at five different levels of abstraction (Dehaene et al., 2015):
i) Transition and timing knowledge
ii) Chunking
iii) Ordinal knowledge
iv) Algebraic patterns
v) Nested tree structures
These levels are represented in Figure 1.3. The first level refers to the knowledge of the transition from one item to the other over time. As for chunking, it refers to the ability to group contiguous items together in single elements, single elements that can be manipulated as units. At a further level of abstraction, sequences can be stored considering the order of each item, so considering what element comes first, what second and so on. At this level (contrary to the transition and timing level), neither the duration of single elements nor the exact timing between elements are encoded. The level of algebraic patterns is the level of abstract categories, in which input regularities are encoded, regardless of the identity of specific items. The levels considered so far cannot account for many processes that we find in one of the more complex type of sequence that we use all day every day, that is language. As pointed out by Noam Chomsky (1956, 1959), complex phrase structures with multiple long-distance dependencies and recursive expressions need hierarchical structures (or

supra-regular grammars, cf. chapter 2), that are represented at the fifth level of Dehaene taxonomy, that of nested tree structures. Nested tree structures are hierarchical, i.e. they show three distinctive properties: i) connectedness: one structure combines all the elements; ii) presence of a root: an element superior to the other elements; iii) no cycles: one element cannot be superior to itself (Fitch and Martins, 2014). More specifically, nested tree structures are hierarchical sequences.



- **TRANSITION AND TIMING**
- **CHUNKING**
  togaba tikibu bekagu togaba bekagu
- **ORDINAL KNOWLEDGE**
  I, II, III, ...
- **ALGEBRAIC PATTERNS**: mental representation of abstract schemata
  A A B A A B A A B
  totoba gogota fofora
- **NESTED TREE STRUCTURES**

Figure 1.3: Sequences Taxonomy, Dehaene et al. (2015)

1.4 Thesis goal

Iversen et al. (2015), discovering that Deaf individuals do synchronize better than hearing individuals to visual flashes, and that hearing individuals synchronize to a bouncing ball as accurate as to an auditory tone, suggested that the ASH does not apply to timing ability in general and may be specifically limited to sequencing behavior. As Dehaene et al. (2015) highlighted, sequencing behavior is a label that need further clarification, because we can store sequences information at at least five different levels of complexity, and actually transition and timing knowledge is only the first level in this taxonomy. So, it has been proposed that the ASH does not apply to the first level of Dehaene taxonomy, but what about the other levels of sequences representation? The main general goal of the present work was indeed to investigate the ASH in light of Dehaene taxonomy and see whether the ASH does apply to some specific sequencing behavior.

Most of my work involved adults participants, and specifically Deaf signers who do not use either hearing aids or CIs. In Conway et al. (2011b), the authors hypothesized that Deaf signers should show sequential processing difficulties because, even if sign languages contain temporal and sequential information, they rely more on simultaneous processing than spoken languages (Wilson and Emmorey, 1997)[2]. On the

---

[2] To be more precise, Conway et al. (2011a) acknowledged that sign languages are *"a rich source of temporal and sequential information"*, but then they seem to neglect this. Even if the issue of sequentiality and simultaneity in sign compared to spoken languages would deserve an entire chapter, here I would like to emphasize just some aspects. Firstly, simultaneous processes occur also in spoken languages, e.g. the disambiguation between declarative sentences and polar questions thanks to prosody in Italian. Secondly, it is true that sign languages phonology presents a characteristic co-occurrence of simultaneous parameters (Brentari, 1998) which is far from the sequentiality of phonemes constituting a spoken word, but i)simultaneous information characterizes also spoken languages phonology (e.g. phonological features) and ii) it is well known that signers make use of mouthings, i.e. the articulation of a word from the local spoken language while signing,

contrary, I believe that having difficulties in processing temporal and sequential information should lead to difficulties in learning a language, regardless of modality [3]. Therefore, my hypothesis is that Deaf signers, as proficient language users, should not show sequence skills impairments. Moreover, I will report a study with deaf children with CI investigating the possibility that timing skills may develop even in a context of delayed exposure to sound and to language.

1.5 Contra Conway et al. (2011b) and the ASH

The validity of Conway et al. (2011b) results, and of the ASH in general, has been recently called into question by Hall et al. (2017) and von Koss Torkildsen et al. (2018).

From a theoretical perspective, Hall et al. contested the validity of drawing inference on the effects of auditory deprivation from a population of deaf children of hearing parents, exposed to sound thanks to CIs. In fact, as I have already discussed, in this population the period of auditory deprivation approximatively overlaps in time with a period of language deprivation. To tease apart the effects of language deprivation and those of auditory deprivation on the development of sequencing skills, the ASH was evaluated by Hall and collaborators by testing a third group of participants: Deaf children without any delay in language exposure, i.e. deaf children of Deaf parents. Hall and colleagues critique was not limited to the theoretical perspective, but to empirical questions as well. Specifically, the authors doubted the replicability of Conway et al. (2011b) results, focusing their discussion on the weak statistical evidence

---

showing therefore the knowledge of some aspect of spoken language sequential phonology (see e.g. Crasborn et al., 2008; Giustolisi et al., 2017). So, it is hard to hypothesize sequence impairments in proficient Deaf sign language users. I might expect differences in sequence processing, but not impairments.

[3] On this point, it is worth notice that people with developmental dyslexia show indeed sequence learning impairments (e.g. Kelly et al., 2002; Howard et al., 2006).

in support of the data. Moreover, Hall and colleagues raised some concerns about the experimental paradigm used in Conway et al. (2011b). In fact, Conway and colleagues testing phase consisted of showing the children a series of sequences one at time, and asking them to reproduce each sequence. Half of the sequences were of the same type of those presented in the familiarization phase, and half were not. The underlying hypothesis was that children should reproduce the sequences of the same type as those presented during familiarization better than unfamiliar sequences. As a matter of fact, children with high working memory span should show no learning effects just because their performance is at ceiling considering both types of sequences (i.e. they can correctly remember familiar sequences as well as non-familiar ones). Therefore, with that paradigm learning effects are detectable only in those children who failed in remembering the presented sequence.

Regardless of the theoretical and empirical concerns, the ASH was challenged to a much greater extent by Hall and colleagues new empirical results. Firstly, they failed to replicate Conway et al. (2011b) results. With three different group of children, i.e. hearing children, deaf children of Deaf parents, and deaf children of hearing parents, Hall et al. found no evidence of learning using Conway's implicit sequential learning task. Secondly, using a Serial Reaction Time Task (SRT Task, Nissen and Bullemer, 1987), all the three groups of children showed learning effects. In the SRT Task, participants give different responses according to the position of a target item. Unbeknownst to participants, item position is determined by fixed transitional probabilities between possible locations. Learning of these fixed transitional probabilities manifests itself in reduced reaction times.

Interestingly, the fact that deaf children of hearing parents also showed learning effects in the SRT Task argues against a possible Language Scaffolding Hypothesis, i.e. that"the development of implicit learning skills may depend less on the temporal and linear structure of *sound* and more on the temporal, hierarchical, and inherently social structure of *language*" (Hall et al., 2017).

Von Koss Torkildsen et al. (2018) used a different sequence learning paradigm, the triplets paradigm (see e.g. Arciuli and Simpson, 2011, 2012, and chapter 3 of the present dissertation) to assess visual sequence learning skills in 34 prelingually deaf children with CI and 34 hearing peers. Results showed that deaf children with CI's performance was comparable to that of hearing children. In addition, in the deaf group the correlation between sequence learning performance and age of implantation or speech perception level was not significant. The Author's discussion focused on the difference between the stimuli used in their task (pictures of aliens) versus those used in Conway et al. (2011b) (colored squares). Von Koss Torkildsen et al. (2018) suspect that differences in verbal rehearsal strategies between deaf children with CI and hearing children might have had a great role in determining Conway et al. (2011b) results (it is likely that participants verbalized squares color to perform Conway et al.'s task, whereas it is unlikely that they could verbalize weird aliens).

All in all, Hall et al. (2017) and von Koss Torkildsen et al. (2018) works shook the validity of the ASH and highlighted the need of addressing this issue in depth, in order to provide enough evidence to eventually disprove a hypothesis that might have severe consequences if taken for granted by clinicians, speech therapists etc..

1.6 Thesis outline

The validity of the ASH and the relationship between sound, language and sequencing abilities is the leitmotif of the present dissertation. The studies that I have collected gave a large contribution in delineating the revised version of the model proposed by Conway et al. 2009 (Figure

1.2), that I will outline in the last chapter of this thesis. As for the Dehaene taxonomy, I will deal with the highest levels, i.e. algebraic patterns and nested tree structures, in chapter 2. Chunking and ordinal knowledge will be dealt with in chapter 3. I will discuss transition and timing knowledge in chapter 4.

Still, every chapter of this dissertation is autonomous, in the sense that every chapter presents specific contents and problems. Chapter 2 is about artificial grammar learning and how Deaf and hearing adults can learn different grammars situated on different levels on the Chomsky hierarchy, a hierarchical representation of complexity of formal grammars, equipped with the computational mechanisms that can accept the languages (i.e. perform grammaticality judgments) produced by these grammars. Chapter 3 presents a visual statistical learning study with Deaf and hearing adults, and focuses on the (different) link between statistical learning performance and reading proficiency in the two groups. Chapter 4 presents a study investigating if deaf children with CI are able to construct the abstract representation of regular isochronous stimuli.

In the last chapter of this thesis, chapter 5, I will draw conclusions in regard to the ASH based on the studies presented in chapters 2, 3, 4 and on the recent works by Iversen et al. (2015); Hall et al. (2017); von Koss Torkildsen et al. (2018).

CHAPTER 2

Artificial grammar learning

In this chapter I deal with algebraic patterns and nested tree structures through a series of artificial grammar learning (AGL) experiments. The experimental work has been carried out in collaboration with Dr. Gesche Westphal-Fitch and Prof. Tecumseh Fitch from the University of Vienna. The first version of the VAGL paradigm was designed during a two weeks period that I spent at the Department of Cognitive Biology of the University of Vienna in June 2015. I took care of data collection in Italy over a ten-months period from July 2015 to May 2016.

2.1 Brief introduction into artificial grammar learning

"Implicit Learning of Artificial Grammars" is the title of Reber's paper in which the AGL paradigm was used for the first time (Reber, 1967). The main reason behind the development of the AGL paradigm was the need to investigate learning mechanisms that allow children to acquire natural languages, mechanisms that were clearly unidentifiable in those proposed by the behaviorist enterprise (in this regard, see Chomsky 1959

contra Skinner 1957). The critical learning mechanisms should have been able to account for children's ability to generate sentences following a given grammar, inferable from the incoming input. Specifically, the AGL paradigm is based on Chomsky (1957) description of language as:

> (...)a set (finite or infinite) of sentences, each finite in length and constructed out of a finite set of elements. (...) the set of "sentences" of some formalized system of mathematics can be considered a language. (p. 13)

From 1967 onwards, Reber's AGL paradigm has become one of the main standard to study sequence learning (e.g. the Reber-like AGL paradigm in Conway et al. 2011b). With slightly variations among studies, the main structure of the paradigm works as follows: There are two phases, exposure and testing. During the exposure phase, participants are exposed to a set of strings derived from an artificial grammar, without being informed of the presence of the grammar. Meanwhile, participants might/might not be involved in a dummy memory task. In the following phase, participants are told that the stimuli are generated by a grammar, and are asked to classify novel strings as grammatical/ungrammatical accordingly. Reber stimuli (i.e. the vocabulary) were visual linguistic stimuli, namely five letters (P,S,T,V,X) and the grammar was a finite state grammar characterized as a Markovian process in which a state-to-state transition produced a letter (Reber 1967, see Figure 2.1). After Reber, AGL abilities in humans have been found with a great variety of different stimuli, from syllable to abstract visual shapes (e.g. Gómez, 2002; Stobbe et al., 2012). Moreover, many AGL experiments demonstrated that not only adults, but also infants can learn abstract rules without explicit teaching (Gomez and Gerken, 1999; Marcus et al., 1999).

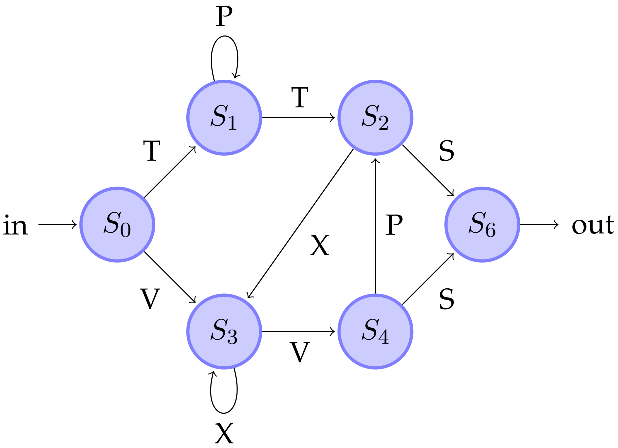


Figure 2.1: Schematic representation of Reber's grammar.
Some examples of grammatical strings: VVS, VXXVPS, TPTXVS, TPPPTXXVS. Strings not beginning with T or V and strings not ending with S would be ungrammatical, as well as a string like: TVXS.

2.2 The Chomsky hierarchy

Formal language theory originated from the attempt to describe human language from a computational perspective (Chomsky, 1956, 1959). Therefore, formal language theory considers language as a theoretically infinite set of strings, with strings being a finite set of symbols coming from a finite alphabet. What type of strings do belong to what language is specified by the language grammar. In this conception, some fundamental components of natural languages are neglected: Firstly, meaning, but also contexts and frequencies of language, and further pragmatics aspects as well.

The Chomsky hierarchy defines four classes of grammars that can generate different type of languages, and four classes of automata (i.e. computational models) that can accept those languages (i.e. perform grammaticality judgments) (see Figure 2.2 and Table 2.1).

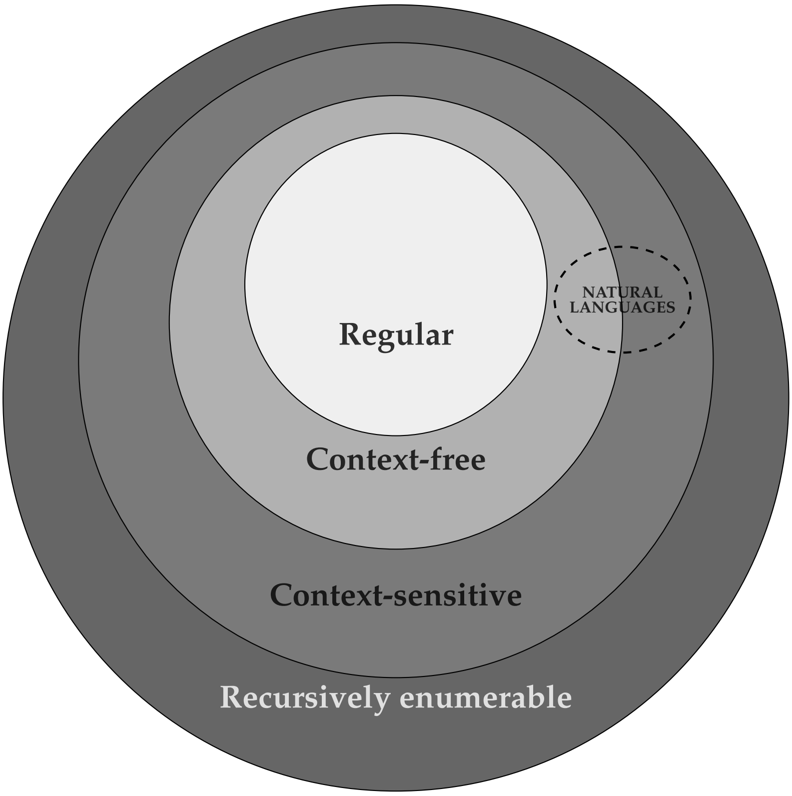


Figure 2.2: Visual representation of the Chomsky hierarchy.
The dotted ellipses indicates natural languages location.

Table 2.1, compiled after Jäger and Rogers (2012), briefly illustrates the form(s) that the rules of each class of grammar take, and some possible language that can be generated.

Table 2.1: The Chomsky hierarchy: grammars generators, rules and possible languages.

Grammar	Rules	Languages
Regular (Finite-state)	A → a A → aB A, B: non-terminal symbols; a: terminal symbol	$(ab)^n$ $ab^n a$
Context-free	A → β A: single non-terminal symbol β: string of symbols	$a^n b^n$ *mirror language*
Context-sensitive	the left-hand side of the rule is never longer than the right-hand side	$a^n b^m c^n d^m$ *copy language*
Recursively enumerable	no restrictions	any language

Regular (finite-state) grammars can be recognized by finite-state automata. Finite-state automata perform grammaticality judgments in a linear fashion, i.e. given an element *n*+1, they can decide if the string is grammatical based on the element *n*. This is because finite-state automata do not have any type of persistent memory available. To recognize higher-level grammars, additional memory systems are needed. Context-free grammars are recognized by push-down automata. Push-down automata have access to a memory system, called *stack*: as they are parsed, new string elements are pushed on a stack. The stored elements can be accessed following the *first-in-last-out* constraint (see e.g. Uddén et al., 2012). Linearly bounded automata can recognize context-sensitive grammars. These type of automata do not follow the *first-in-last-out* constraint. Still, they cannot perform operation over a tape of infinite length. Only Turing machines, the highest level of automata, can recognize any language with no restrictions (see Jäger and Rogers, 2012, for a detailed discussion).

One of Chomsky's major interest (and actually, not just his interest) was to locate natural languages on this hierarchy. It was extremely clear to Chomsky (1957) that regular grammars cannot account for all the complex phenomena occurring in natural languages. Regular grammars produce finite-state languages: As I have mentioned before, the Reber's grammar is an example of regular grammar. From the first symbol (initial state) to the last one (final state, each symbol is generated from a state), the process that produces a finite-state string can be described with the path that departs from each state, and with the transitional probabilities from one state to the other (see again Figure 2.1. States are represented with the capital letter S and paths with arrows). Please refer to Chomsky for a detailed explanation of why English (and as extension, natural languages) is not a finite-state language. The main argument lies in the presence, in the English language, of a (theoretically) unlimited number of nested dependencies. I will briefly summarize this argument.

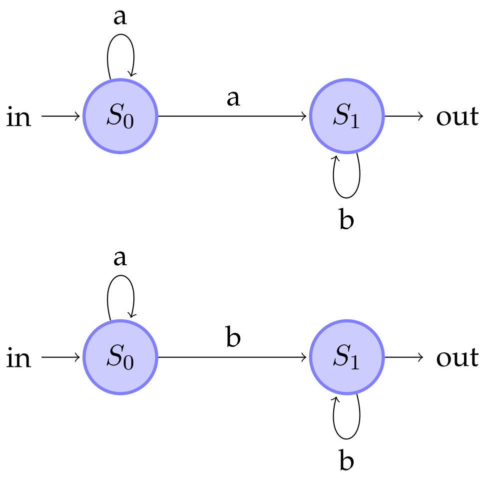
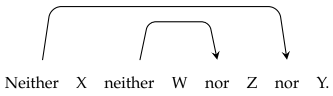


Figure 2.3: Possible grammars for a $a^n b^m$ language.

Let's compare a language like $a^n b^m$, composed by any number of *a*s followed by any number of *b*s, and a language like $a^n b^n$, with any number of *a*s followed by the exact same number of *b*s. $A^n b^m$ is a regular language (Figure 2.3), whereas $a^n b^n$ is not, because it requires to count and compare the number of *a* with the number of *b*. To describe $a^n b^n$ we need a hierarchical phrase structure, made for example from the following grammar (which is not the only possible grammar, see Jäger and Rogers 2012):
S → aSb;
S → ab.
Natural languages have $a^n b^n$ strings, like the neither/nor construction in English: neither X nor Y, which is reiterable (Neither X neither W nor Z nor Y, see Figure 2.4) an in principle unlimited number of times.



***Neither** did Mary state that she **neither** drinks **nor** smokes **nor** did anybody believe it.*

Figure 2.4: Context-free structure in English

Regular grammars cannot generate languages with an unbounded number of nested dependencies, like the structure presented in Figure 2.4, therefore English is not a finite-state language. A further explicative example of a context-free structure in English is the following: *"The woman that the dogs are chasing is screaming"* , consisting of a center embedded object relative clause.

Moreover, Chomsky pointed out that context-free grammars cannot describe all the processes encountered in natural languages, but this second idea was not so broadly accepted, and required further work and clarifications. As reviewed in Jäger and Rogers (2012), the demonstration that natural languages are not context-free languages arrived thanks to the simultaneous, albeit independent work of Huybregts (1984), Shieber (1985) and Culy (1985), who proved that neither Swiss-German (Huybregts and Shieber) nor Bambara – the national language of Mali – (Culy) are context-free languages. The example reported in Jäger and Rogers, from Shieber (1985), is depicted in Figure 2.5. The Swiss-German sentence *Dass mer d'chind em Hans es Huus lönd hälfe aanstriiche* means "that we let the children help Hans paint the house".

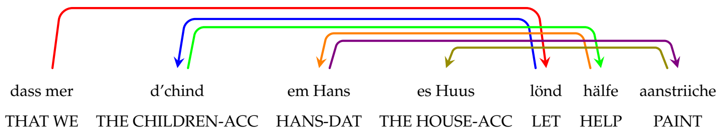


Figure 2.5: Non-context-free structure in Swiss-German: Crossed dependencies.

Context-free grammars can generate languages with an unbounded number of nested dependencies, but they cannot generate languages with (an unbounded number of) crossed dependecies, as the one of Figure 2.5. Subsequently, the Chomsky hierarchy has been refined by the addition of an intermediate layer between context-free grammars and context-sensitive grammars: the layer of mildy context-sensitive languages (Joshi et al., 1991). To date, computational linguists agree that all processes encountered in natural languages can be generated by mildy-context-sensitive grammars.

2.3 Formal language theory and AGL

As reviewed in Fitch et al. (2012), Chomsky's work contributed to the development of Miller's project Grammarama (Miller, 1958, 1967). Born with the purpose of investigating human ability to acquire regular languages, Miller's project ended with the proposal of the supra-regular hypothesis, according to which we, i.e. humans, might favor learning of supra-regular grammars (in the sense of non-regular grammars), because they better reflect our cognitive approach to learning rules (Fitch et al., 2012). After Miller and Chomsky's joint work, human rule learning has been extensively investigated using the AGL paradigm. The main focus was to compare implicit versus explicit learning (Perruchet and Pacteau, 1990; Mathews, 1990). On the contrary, there was no interest on the grammars *per se* until the work of Fitch and Hauser (2004). Using two auditory presented grammars, a finite-state $(ab)^n$ grammar and a context-free $a^n b^n$ grammar, Fitch and Hauser showed that humans could easily master both of them. On the contrary, cotton-top tamarins acquired the finite-state grammar, but not the context-free grammar. The Authors concluded their work with the following hypothesis:

> If nonhumans are "stuck" trying to interpret PSG-generated stimuli at the FSG level, it would make PSG stimuli seem much more complex to them and perhaps even unlearnable in finite time [PSG = phrase state grammar, i.e. context-free grammar; FSG = finite-state grammar]. Though the evolution of well-developed hierarchical processing abilities in humans might have benefited many aspects of cognition (e.g., spatial navigation, tool use, or social cognition), this capability is one of the crucial requirements for mastering any hu-

> man language. (p.380)

To date, the idea that dealing with supra-regular grammars might be a distinctive feature of humans, albeit challenged by some studies (see Gentner et al. 2006 but also Van Heijningen et al. 2009 and Zuidema 2012 for criticisms and ten Cate and Okanoya 2012 for a review on AGL in nonhuman species), continues to be plausible. Nevertheless, the debate is still open.

2.3.1 My interest into AGL

My approach to AGL was guided by Stobbe et al. (2012) paper on visual AGL (VAGL) in kea, pigeons and humans. Using the same grammars of Fitch and Hauser (2004), i.e. the regular $(ab)^n$ grammar and the supra-regular $a^n b^n$ grammar, the Authors showed that, contrary to humans, kea and pigeons could not learn supra-regular rules. Crucially, the grammar were presented in the visual modality. A and B categories were abstract visual patterns and all grammar strings were presented simultaneously to minimize working memory load (this was done to allow a more reliable comparison between humans and birds). Those results were consistent with previous studies considering other species and presenting the grammars in the auditory modality. Moreover, those results brought further evidence to the fact that humans can learn complex sequences regardless of modality. Still, it wasn't clear how far humans can go in learning supra-regular grammars, especially considering the visual modality and how different working memory load might impact humans learning ability. Those questions, and, as I will explain in the next session, the visual nature of Stobbe et al. (2012) AGL paradigm, led me to Study 1.

2.4 Study 1: visual AGL in deaf and hearing adults

The main goal of the present study was to explore the ability of human adults to acquire artificial grammars of differing complexity on the Chomsky hierarchy. Even if computationally speaking nested dependencies (found in context-free grammars) are easier to process than crossed dependencies (found in context-sensitive grammars), it seems that humans prefer the latter. It is well documented that Dutch has a preference for crossed dependencies between verbs and arguments, while in German nested dependencies are more common. Starting from this evidence, Bach et al. (1986) investigated how Dutch and German speakers process crossing and nested dependecies, respectively. They found an advantage for Dutch speakers, suggesting than crossed dependencies are somehow ”easier” than nested dependecies. The Authors interpreted these data as proving the push-down stack model of sentence processing wrong (see Figure 2.6. If this was the right model to represent the processing of nested and crossed dependecies, the working memory load to process crossed dependencies would be higher than that needed to process nested dependencies, making Bach et al. results hard to explain). As a consequence, Joshi (1990) introduced embedded pushdown automata, which use multiple embedded stacks allowing the processing of crossed dependencies. Embedded pushdown automata can also process nested dependencies, but they are computationally more costly than crossed dependencies in terms of numbers of items that have to be stored during computation. This potentially explains why crossed dependencies, despite being formally more complex, may nonetheless be processed more easily, since load on working memory may be less for crossed than for nested dependencies.

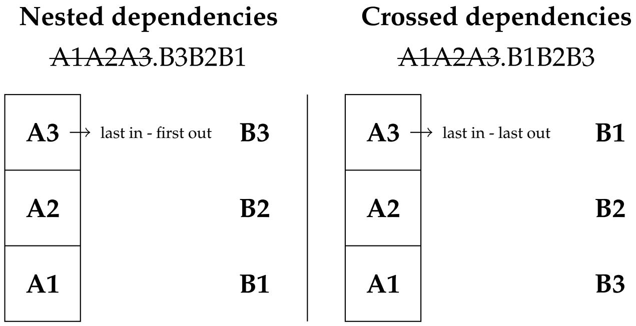


Figure 2.6: Push-down stack model of sentence processing.

When A1, A2 and A3 are parsed, only the top element A3 is available to be compared with the first B element. So, in nested dependecies the element A3 is the last to enter the memory system and the first to leave as B3 is processed. In crossed dependencies, the element A3 is the last to enter the memory system and the last to leave. To process crossed dependencies, a powerful memory system that can reach A1 keeping A2 and A3 in memory is needed, i.e. extra stacks.

Another important factor that has to be considered is the presence of long distance dependencies versus adjacent dependecies, with the former more difficult to be processed and learned (see Uddén et al. 2012 for an AGL experiment on long distant dependencies and Chesi and Moro 2014 for a review and discussion on the complexity of long distance dependencies processing).

In the present study, we constructed different AGL experiments in the visual modality. The grammars we used had at least one long-distance dependency and they were located on different levels on the Chomsky hierarchy. Specifically, we employed a finite-state grammar (ab^na), a context-free grammar (mirror grammar, ww^r) and a (mildly) context-sensitive grammar (Copy grammar, ww). The ab^na grammar generates strings beginning and ending with an A element, with a variable number of B elements in the middle (e.g. A.BB.A; A.BBBBB.A). As I said, it is a regular grammar and a possible generating machine is represented in Figure 2.7.

The ww^r grammar generates strings in which the second half is specular to the first one (e.g. AAB.BAA; ABAA.AABA), whereas the ww grammar generates strings in which the second half is the reduplication of the first one (e.g. AAB.AAB ABAA.ABAA). The mirror grammar is context-free, and a push-down memory system is required in order to recognize this grammar. To recognize the copy grammar, a more powerful memory system is needed, with no *first-in-last-out* constraint (see again Figure 2.6).

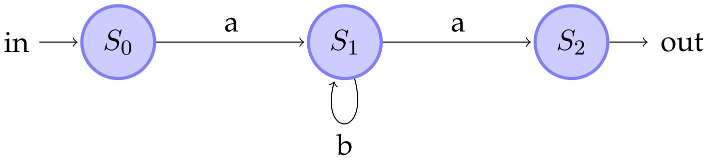


Figure 2.7: ab^na grammar

VAGL and the auditory scaffolding hypothesis The fact that grammars were presented in the visual modality ensured the possibility to use the same material with hearing and deaf participants. In fact, my peculiar interest was to investigate if deaf people, and specifically Deaf adult signers, could perform an AGL task in a similar way to hearing adults. Learning of the finite state grammar should prove the ability to extract algebraic patterns from an incoming sequential input, whereas learning of the mirror and the copy grammar should be an evidence for the ability to learn nested tree structures and computationally more complex structures. The task, no matter which grammar is considered, would be virtually undoable for deaf individuals if the auditory scaffolding hypothesis (ASH, Conway et al., 2009) held true. As I explained in the previous chapter, according to the ASH the development of sequential skills is sustained by hearing experience. Therefore, deaf individuals should show impairments in tasks targeting sequential abilities. On the contrary, it could also be the case that Deaf people, using a language conveyed and perceived in the visual modality, might be better than hearing people in learning complex grammars *presented visually*.

Stimuli Stimuli were videos in mp4 showing visual strings. Strings were composed by tiles, colorful abstract decorated small squares. Tiles belonged to one of two distinct categories, that we called A and B. A tiles were decorated with rounded, nested grey/purple shapes, whereas B tiles were decorated with un-nested reddish and greenish angular shapes. Tiles were like the ones used in Stobbe et al. (2012) (see Figure 2.8). To build strings, tiles were randomly chosen without replacement.

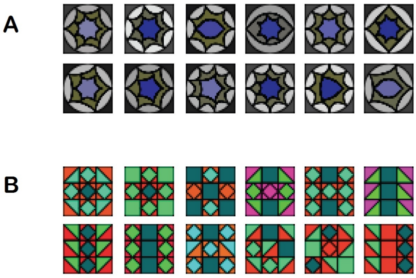


Figure 2.8: Example of A and B tiles.

Ethics The studies were approved by the ethics committee of the University of Milano-Bicocca.They were carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki).

Experiments Study 1 is composed by four experiments. Experiment 1-3 could be considered as pilot experiments for the designing of experiment

4.

2.4.1 Experiment 1 - transient modality

Methods

Participants The final sample of participants consisted of 20 hearing students (5 males, mean age=20.9, SD=3.0) from the University of Milano-Bicocca. Nine additional participants were excluded because they could not learn the warm-up grammar (see below). All participants had normal or correct to normal vision. Participants gave their written informed consent prior to taking part and received course credits.

Procedure Prior to testing, participants were trained with an $(ab)^n$ warm-up grammar. They were tested with a 15-trials session, and they had to reach a criterion of 12/15 corrects (exact binomial test, p = .02) to proceed with the experiment. If they failed after three training runs, they were excluded from further testing. After exposure, participants received the following oral instructions: *Ora il tuo compito è decidere se ogni nuovo video che vedi segue lo stesso schema di quelli della fase di familiarizzazione oppure no*, that could be translated as: Now your task is to decide if each new video that you see follows the same schema as those of the familiarization phase or not. Each participant was tested on each grammar in random order. The exposure phase for each grammar lasted about two minutes. During the exposure phase participants saw 30 grammatical strings with N=2, N=3[1]. All participants saw the same strings, but in a different randomized order. The presentation modality was the following: Each symbol appeared on a black background at a rate of 6 frames per second (i.e. a new element every 166 ms) directly adjacent to the location

---

[1] Considering ww and ww^r grammars, N equals half of the number of symbols contained in a given string. As for the ab^na grammar, N equals n.

of the previous element. Each symbol disappeared right after it was presented. In this way, the whole string was never displayed on the screen. We called this modality transient modality. In the middle of the sequence (mirror and copy grammar) or between As and Bs (ab^na) we presented a black rectangle (16 x 20 pixels) for 166 ms, which determined a spatial and temporal gap between "phrases", highlighting string structure.

In the test phase individual strings were showed. After the final symbol had appeared, the screen went blank and participants had to perform the grammaticality decision by pressing a labeled green SI (yes) or red NO key on a keyboard. Response time was not limited and there was no feedback. Answers given before the string was completed were not recorded. The testing phase was composed by 75 stimuli. There were two types of grammatical stimuli: i) Grammatical (20 novel exemplars of N=2 and N=3), and ii) Grammatical extensions (10 strings of N=4, that is grammatical strings of a N never shown during exposure), and two types of ungrammatical foils: i) Foil missing tile (20 strings with missing elements, i.e. incomplete dependencies. 10 with N=2,3 and 10 with N=4), and ii) Foil wrong tile (25 strings with a wrong element, e.g. ABBABA. 15 with N=2,3 and 10 with N=4). I will refer to these categories as type of strings.

Results and analysis

Overall, participants responded correctly to 87% of grammatical stimuli considering the ab^na grammar, 74% of grammatical stimuli in the ww^r grammar and 65% of grammatical stimuli in the ww grammar. As for ungrammatical stimuli, accuracy was 86% for the ab^na grammar, 55% for the ww^r grammar and 58% for the ww grammar. Table 2.2 clarifies how accuracy varied between type of strings.

Table 2.2: Transient modality: group results

Grammar	Type of string	Mean accuracy (SD)
ab^na	Grammatical (N=2,3)	94 (24)
	Grammatical extension (N=4)	72 (45)
	Foil missing tile	97 (18)
	Foil wrong tile	78 (41)
ww^r	Grammatical (N=2,3)	78 (42)
	Grammatical extension (N=4)	66 (48)
	Foil missing tile	61 (49)
	Foil wrong tile	50 (50)
ww	Grammatical (N=2,3)	84 (37)
	Grammatical extension (N=4)	27 (45)
	Foil missing tile	56 (50)
	Foil wrong tile	60 (49)

We considered individual participant performance on type of strings, evaluated by the following criteria: grammatical→ at least 15/20 trials correct (exact binomial test p=.02); grammatical extension→ at least 9/10 trials correct (exact binomial test p=.01); foil missing tile→ at least 15/20 trials correct (exact binomial test p=.02); foil wrong tile→ at least 18/25 trials correct (exact binomial test p=.02).

Table 2.3: Transient modality: individual performance. Participants= 20. Cells report the number of participants who performed above chance.

Grammar	Grammatical	Grammatical extension	Foil missing tile	Foil wrong tile
ab^na	19	10	20	15
ww^r	12	8	2	4
ww	15	1	3	5

As Table 2.3 shows, half of the participants could generalize the rule for the finite state grammar (see the Grammatical extension column). However, half of them could not. Considering the finite state grammar, the vast majority of participants correctly rejected ungrammatical strings. The performance was worst in the two supra-regular grammars: In these grammars, performance was poor considering grammatical and ungrammatical strings.

To sum up, for all grammars participants were good at recognizing grammatical patterns of N=2 and N=3, but most of them could not reject ungrammatical strings and accept grammatical extensions. This seems to indicate that the vast majority of participants did not learn the rule.

2.4.2 Experiment 2 - typewriter modality

The poor performance reported in 2.4.1 led us to design a different presentation modality, that we called "typewriter modality", as I will explain shortly. The purpose of the new modality was to decrease working memory load, which might result in better performance.

Methods

Participants Twenty hearing students (6 males, mean age=23.9 years, SD=3.4) from the University of Milano-Bicocca participated. One additional participant was excluded as he did not reach the warm-up criterion. All participants had normal or correct to normal vision. Participants gave their written informed consent prior to taking part and received course credits.

Procedure The procedure was the same as for transient modality (2.4.1), but with a remarkable difference in how strings were presented: Each symbol remained on the screen until the whole string was completed. We called this modality typewriter modality because the entire sequence was completed one by one, akin to a word being typed: this allowed us to minimize working memory load.

Results and analysis

Overall, participants responded correctly to 87% of grammatical stimuli considering the ab^na grammar, 77% of grammatical stimuli in the ww^r grammar and 75% of grammatical stimuli in the ww grammar. As for ungrammatical stimuli, accuracy was 94% in the ab^na grammar, 88% in the ww^r grammar and 93% in the ww grammar. As Table 2.4 shows, participants were overall very accurate in detecting grammatical strings of N=2 and N=3. It seems also that participants could satisfactorily reject ungrammatical stimuli. However, it is quite clear that participants could not generalize the rule and accept grammatical extensions.

Table 2.4: Typewriter modality: group results

Grammar	Type of string	Mean accuracy (SD)
ab^na	Grammatical (N=2,3)	99 (12)
	Grammatical extension (N=4)	65 (48)
	Foil missing tile	97 (17)
	Foil wrong tile	92 (27)
ww^r	Grammatical (N=2,3)	95 (22)
	Grammatical extension (N=4)	40 (49)
	Foil missing tile	98 (15)
	Foil wrong tile	80 (40)
ww	Grammatical (N=2,3)	97 (17)
	Grammatical extension (N=4)	31 (46)
	Foil missing tile	99 (10)
	Foil wrong tile	88 (33)

We investigated individual participant performance, evaluated by the same criteria as in 2.4.1. Individual performance analysis confirmed what observed at the group level: all the participants could recognize grammatical strings of N=2 and N=3, but only half of them could accept grammatical strings of a new N considering the finite state grammar, whereas the vast majority of participants could not generalize to N=4 in the two supra-regular grammars. Participants also rejected ungrammatical strings for all three grammars. This pattern suggests that participants learned a subset of the intended rule, covering only N=2 and N=3. By this rule, participants did not accept the foils, but also rejected grammatical strings that went beyond the string length they had been exposed to in the first phase.

Table 2.5: Typewriter modality: individual performance. Participants= 20. Cells report the number of participants who performed above chance.

Grammar	Grammatical	Grammatical extension	Foil missing tile	Foil wrong tile
ab^na	20	10	19	18
ww^r	20	3	20	15
ww	20	5	20	17

2.4.3 Modalities comparison

Analysis

To compare results in the two modalities, we performed a mixed effects logistic regression analysis with the statistical software R (R Core Team, 2016). The dependent variable was a binary outcome variable: accuracy. We entered in the model two categorical predictors, modality (two levels: transient vs. typewriter), grammar (three levels: mirror grammar vs. finite state grammar and copy grammar) and their interaction. Results are reported in Table 2.6: The interaction was significant, and accuracy was higher in the typewriter compared to the transient modality.

Even if results were better in the typewriter modality, and participant were able to accept grammatical strings of N=2 and N=3 and reject un-

Table 2.6: Modality comparison: logistic regression analysis

Fixed effects	β	SE	z value	
finite state grammar	1.60	0.10	15.827	***
copy grammar	-0.08	0.08	-0.96	
modality	1.47	0.26	5.742	***
fin. state grammar:mod	-0.67	0.16	-4.173	***
copy grammar:mod	0.28	0.14	2.030	*

* p<.05; *** p<.001
Reference levels: modality = transient;
grammar = mirror grammar

grammatical foils, we were not satisfied by these results either, because participants could not generalize the rule and accept grammatical strings of N=4 (see Table 2.5). This led us to the design of experiment 4, reported in 2.4.7.

2.4.4 Experiment 3 - Deaf participants

To verify the possibility that signers, as users of a language visually perceived, might perform better than non-signers in an AGL task presented in the visual modality, the same task as in experiment 2 was administered to a group of Deaf proficient signers. Experiment 2 and experiment 3 were performed during the same period.

Methods

Participants Eleven Deaf member of the Deaf Institute of Turin participated (5 males, mean age=24.8 years, SD=3.9). With 2 exceptions (2 M), they were all university students. Four additional participants were excluded as they did not reach the warm-up criterion. Participants were all born deaf and none of them had any associated disability or further sensory deficits. All participants were fluent LIS signers, and they all considered LIS as their primary means of communication. Seven of them were native signers, being exposed to LIS from birth (N=6) or from age 1. One participant was first exposed to LIS at age 6, and three participants were first exposed to LIS later, during adolescence. All participants had normal or correct to normal vision. Participants gave their written informed consent prior to taking part and received a monetary reimbursement for their participation.

Procedure As for typewriting modalities with hearing participants (2.4.2), but with the difference that instructions were given in LIS through a videotaped message.

Results

Considering the finite state grammar, participants responded correctly to 74% of grammatical and 76% of ungrammatical stimuli. As for the mirror grammar, participants responded correctly to 75% of grammatical and 61% of ungrammatical stimuli. Finally, participants responded correctly to 72% of grammatical and 67% of ungrammatical stimuli in the copy grammar. Detailed results are reported in Table 2.7

Individual participant performance on strings in the test phase was evaluated following the same criteria used for the hearing participants: grammatical→ at least 15/20 trials correct (exact binomial test p=.02); grammatical extension→ at least 9/10 trials correct (exact binomial test p=.01); foil missing tile→ at least 15/20 trials correct (exact binomial test p=.02); foil wrong tile→ at least 18/25 trials correct (exact binomial test p=.02).

Most of the Deaf participants performed above chance considering grammatical stimuli of N=2 and N=3, but they also rejected grammatical extensions. Moreover, they could reject ungrammatical foils with a miss-

Table 2.7: Deaf participants: group results

Grammar	Type of string	Mean accuracy (SD)
ab^na	Grammatical (N=2,3)	85 (36)
	Grammatical extension (N=4)	52 (50)
	Foil missing tile	87 (33)
	Foil wrong tile	67 (47)
ww^r	Grammatical (N=2,3)	89 (32)
	Grammatical extension (N=4)	49 (50)
	Foil missing tile	83 (38)
	Foil wrong tile	44 (45)
ww	Grammatical (N=2,3)	91 (28)
	Grammatical extension (N=4)	35 (48)
	Foil missing tile	85 (36)
	Foil wrong tile	53(50)

Table 2.8: Typewriter modality, Deaf participants: individual performance. Participants= 11. Cells report the number of participants who performed above chance.

Grammar	Grammatical	Grammatical extension	Foil missing tile	Foil wrong tile
ab^na	8	2	9	5
ww^r	9	3	7	2
ww	10	2	7	4

ing tile, but the vast majority could not reject ungrammatical foils with a wrong tile. These results suggests that many Deaf participants counted the number of stimuli in the exposure phase and learned a rule anchored to the number of stimuli presented.

Being a possible confounding, we checked whether performance of the three late signers was worse than performance of the other Deaf participants. This was not the case, on the contrary the situation appeared to be extremely variable, and this was visible considering the three late signers only. One late signer was above chance considering all grammars and string types. On the other hand, one late signer performed at chance considering all grammars and string types. The third late signer performed well on the mirror and the copy grammar, but not in the (usually easier) finite state grammar.

2.4.5 Typewriter modality: Deaf vs. hearing participants

Participants We compared group performance of the Deaf participants with group performance of the hearing participants in the typewriter modality. The two groups were matched in age (t=-0.68, p=.49). With the exception of two Deaf participants, all participants were university student. An important difference is that all hearing participants were psychology students, whereas the Deaf participants had different type of majors.

Analysis

Figure 2.9 shows that accuracy was overall higher for the hearing participants, with the exception of grammatical stimuli in the mirror and copy grammar. The comparison of Table 2.4 and Table 2.7 and of Table 2.5 and Table 2.8 shows how both groups have indeed poor performance with supra-regular grammars grammatical extensions. To compare Deaf and hearing performances, we performed a mixed effects logistic regression analysis with the statistical software R (R Core Team, 2016).

Grammatical stimuli The dependent variable was accuracy, dichotomically coded. We entered in the model three categorical predictors: Group (two levels: Deaf and hearing), grammar (three levels: Finite State Grammar, Mirror Grammar and Copy Grammar. Reference level: Finite State Grammar) and type of stimulus (two levels: correct N=2, N=3 and correct extensions, N=4). We added random intercepts for each participant

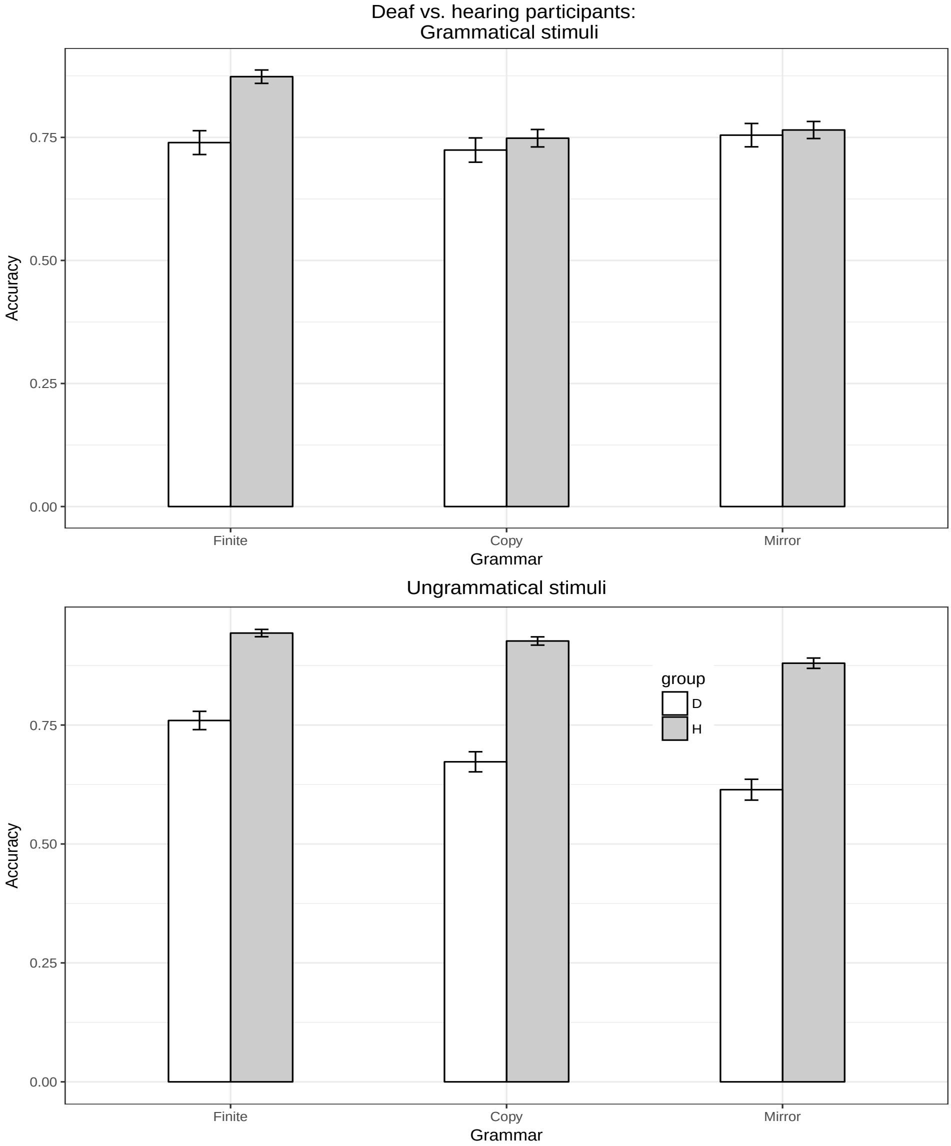


Figure 2.9: Deaf (white bars) and hearing (gray bars) participants: mean accuracy on grammatical and ungrammatical stimuli.

Accuracy is plotted on the vertical axis, and grammar on the horizontal axis. Bars indicate the standard error.

Table 2.9: Deaf vs. hearing participants - Grammatical stimuli: logistic regression analysis.

Fixed effects	β	SE	z value	
group	3.25	0.61	5.369	***
grammar ww	-0.11	0.21	-0.525	
grammar wwr	0.11	0.21	0.533	
type	-2.46	0.18	-13.768	***
group:grammar ww	-1.68	0.32	-5.320	***
group: grammar wwr	-1.68	0.32	-5.333	***
group: type	-2.18	0.30	-7.338	***

*** $p<.001$
Reference levels: grammar = finite state grammar;
group = deaf;
type = correct N=2, N=3

and for each item. We tested a model with the three main effects and two two-way interactions, between group and grammar and between group and type of stimulus. Results are reported in Table2.9. The group * grammar interaction is depicted in Figure 2.10: In the hearing group accuracy is higher in the finite state grammar compared to the two supra-regular grammars, whereas this is not the case for the deaf group. The group * type interaction is depicted in Figure 2.11: Hearing participants performed better than deaf participants considering grammatical stimuli of N=2 and N=3, whereas both groups performed equally poor with grammatical extensions.

Ungrammatical stimuli The dependent variable was accuracy, dichotomically coded. We entered in the model three categorical predictors: Group (two levels: Deaf and hearing), grammar (three levels: Finite State Grammar, Mirror Grammar and Copy Grammar. Reference level: Finite State

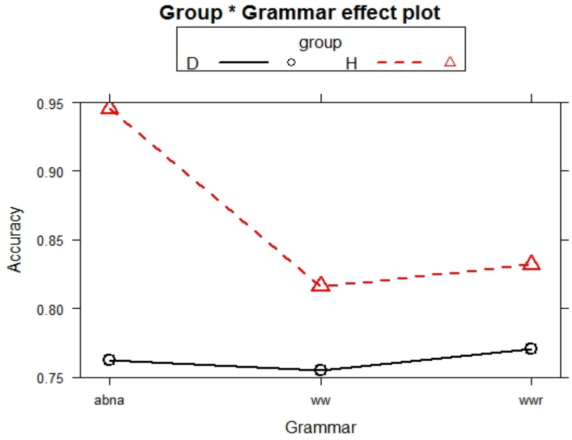


Figure 2.10: Deaf vs. hearing participants - Grammatical stimuli: Group * grammar interaction.
Accuracy is plotted on the vertical axis and grammar on the horizontal axis. Deaf participants are represented with circles and hearing participants with triangles.

Grammar) and type of stimulus (two levels: missing tile and wrong tile). We added random intercepts for each participant and for each item. We tested a model with the three main effects and two two-way interactions, between group and grammar and between group and type of stimulus. The group * grammar interaction was not significant, so we dropped without decreasing the model's goodness of fit (χ^2 = 4.76, p=.09). The reduced model is reported in Table2.10. The group * type of stimulus interaction is depicted in Figure 2.12. Overall, hearing participants performed better than Deaf participants, however, the difference was higher considering foils with a missing tile.

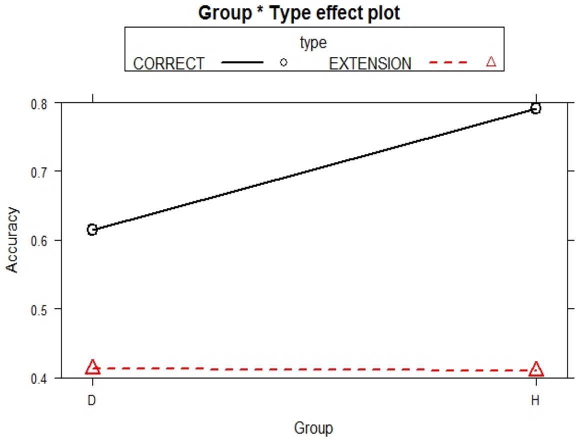


Figure 2.11: Deaf vs. hearing participants - Grammatical stimuli: Group * type of stimulus interaction.

Accuracy is plotted on the vertical axis and group on the horizontal axis. Grammatical stimuli of N=2 and N=3 are represented with circles and grammatical extensions with triangles.

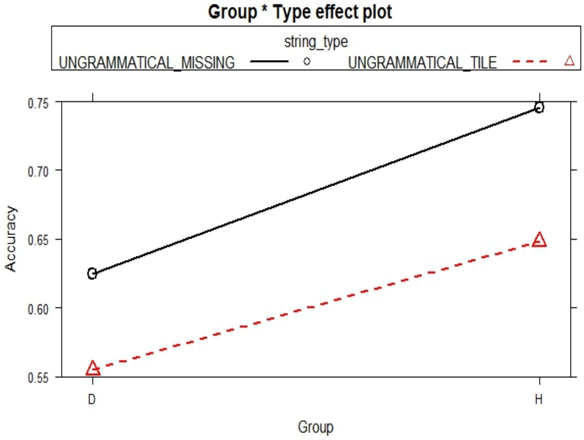


Figure 2.12: Deaf vs. hearing participants - Ungrammatical stimuli: Group * type of stimulus interaction.
Accuracy is plotted on the vertical axis and group on the horizontal axis.
Ungrammatical stimuli with a missing tile are represented with circles and ungrammatical stimuli with a wrong tile with triangles.

Table 2.10: Deaf vs. hearing participants - Grammatical stimuli: logistic regression analysis.

Fixed effects	β	SE	z value	
group	3.61	0.82	4.384	***
grammar ww	-0.62	0.14	-4.357	***
grammar wwr	-1.20	0.14	-8.533	***
type	-2.10	0.18	-11.387	***
group: type	-0.80	0.29	-2.794	**

** p <.01; *** p<.001
Reference levels: grammar = finite state grammar;
group = deaf;
type = missing tile

To sum up Deaf signers, proficient user of a visual language, did not perform better than hearing non signers in this VAGL task. On the contrary, they performed overall worse.

2.4.6 Experiment 1, 2 & 3 - interim summary

In experiment 1, 2 and 3 we used a VAGL paradigm to test the human ability to acquire rules at different levels on the Chomsky hierarchy. We tried two different modalities of stimuli presentation, the transient modality, more demanding for working memory because strings were never presented entirely, but one element at time and the typewriter modality, where strings were progressively displayed, one element next to the other until the whole string was completed, being therefore less demanding for working memory. In the transient modality, psychology students could recognize previously presented patterns, but they could not acquire the underlying rules. In the typewriter modality, the rule was easily acquired for the finite state grammar, but many participants showed an incomplete learning in the two supra-regular grammars. We tested a small sample of Deaf participants with the typewriter modality only. The Deaf participants showed good pattern recognition abilities, but they could not generalize the rules, showing also no advantage for the finite state grammar over the supra-regular grammars. Crucially, using a language in a visual modality does not seem to facilitate a VAGL task. Still, the Deaf-hearing comparison had some limitations because the Deaf group was less used to taking part into experimental studies and the sample size was smaller than for the hearing group.

2.4.7 Experiment 4

As reported before, experiment 1, 2 & 3 could be considered pilot studies for the design of experiment 4. We knew that performance in the typewriter modality was better than in the transient modality, but still most of the participants could not generalize the rule considering the two supra-regular grammars. To improve generalization, we decided to introduce a few sequences of N=5 in the familiarization phase. The rationale behind this was to make participants understand that it was not important for sequences to be exactly of N=2 or N=3 in order to be considered grammatical. To test for generalization, we maintained the N=4 sequences in the testing phase, and to test for generalization beyond the previous encountered N, we added N=6 sequences as well. Moreover, all participants performed a visuo-spatial working memory task. It was our purpose to see if the VAGL task was more difficult for people with lower visuo-spatial span.

Methods

Participants One group of Deaf people and one group of hearing people participated. Deaf participants were 15 LIS signers (mean age = 33 years; SD= 14 years; range = 18 - 62; 7 Females, 8 Males) recruited from the members of several Italian Deaf Institutes (Milan, Monza, Turin, and Verona). They were all born deaf and none of them had any associated disability or further sensory deficits. Seven out of fifteen Deaf participants (47%) were native signers, being exposed to LIS from birth; eight out of fifteen (53%) were first exposed to LIS during childhood. At the moment of testing, all participants were fluent LIS signers and used LIS as everyday means of communication. The mean number of years of education was 13.7 years (SD=2.6). Hearing participants were 15 hearing Italian speakers (7 Females, 8 Males) with no knowledge of LIS or any other sign language. They were recruited from the Milan metropolitan area through on-line social media and flyers. They were matched with the Deaf participants in age (mean age= 34 years; SD= 15 years; range = 18 - 59; t = -0.16, p=.87) and level of education (mean years of education= 13.5 years; SD= 2.4; t= -0.07, p=.95). Overall cognitive abilities of the two groups of participants were assessed using Raven's Colored Progressive Matrices (Raven, 1965). Raw scores were corrected following Basso et al. (1987). Mean corrected scores for the Deaf participants was 31.67 (SD= 4.76), for the hearing participants was 33.33 (SD= 3.11), with no difference between groups (t=1.14, p=.27). All participants had normal or correct to normal vision. Participants gave their written informed consent prior to taking part to the experiment and received €20 reimbursement for their participation.

All participants performed a visuo-spatial working memory task: the Corsi-Block tapping task (Corsi and Michael, 1972). The task was administered using the nine square blocks positioned on a plastic board. The two groups obtained similar results: the mean Corsi score for the Deaf participants was 5.67 (SD=0.90) and for the hearing participants was 5.53 (SD=0.99). The difference was not significant (t= 0.38, p=.70).

Materials and Procedure All participants received instruction in their first language, i.e. Deaf participants received all instructions in LIS and hearing participants in Italian. Each participant was tested on the three target grammars in a random order. For each grammar, the procedure was divided in two phases: exposure and testing. In each phase, all participants saw the same sequences, but in a different randomized order. During exposure (duration approx. 2 minutes), participants saw 30 grammatical sequences with N=2, 3 and 5 . During testing, participants saw 87 individual strings: 36 grammatical (N=2 and N=3, and extensions of N=4 and N=6) and 51 ungrammatical (again N=2, 3, 4 and 6). Ungrammatical strings could be sequences with a missing element or sequences with the correct number of elements, but incorrect category members (see Appendix A for specific information about stimuli and for examples of grammatical and ungrammatical sequences). Participants' task was to indicate whether each sequence followed the same schema as those seen during the exposure phase or not. They did that by pressing a yes/no key on a keyboard. Response time was not limited and no feedback was given. The experimental session was preceded by a training session during which participants were exposed to the warm-up finite state grammar, $(ab)^n$. Success on this grammar (accuracy >chance level, i.e. accuracy >12/15, Exact binomial test, p = .02) was the prerequisite for accessing the experimental session.

Results and analysis

Deaf participants Considering the finite state grammar, participants responded correctly to 81% of grammatical and 83% of ungrammatical stimuli. As for the mirror grammar, participants responded correctly to 82% of grammatical and 65% of ungrammatical stimuli. Finally, participants responded correctly to 81% of grammatical and 66% of ungrammatical stimuli in the copy grammar. Detailed results are reported in

Table 2.11

Table 2.11: Deaf participants: group results

Grammar	Type of string	Mean accuracy (SD)
ab^na	Grammatical (N=2,3)	92 (27)
	Grammatical extension (N=4,6)	68 (47)
	Foil missing tile	92 (29)
	Foil wrong tile	75 (43)
ww^r	Grammatical (N=2,3)	93 (26)
	Grammatical extension (N=4,6)	69 (46)
	Foil missing tile	75 (43)
	Foil wrong tile	55 (50)
ww	Grammatical (N=2,3)	95 (22)
	Grammatical extension (N=4,6)	65 (48)
	Foil missing tile	76 (43)
	Foil wrong tile	56 (50)

Hearing participants In the finite state grammar, participants responded correctly to 95% of grammatical and 95% of ungrammatical stimuli. Considering the the mirror grammar, participants responded correctly to 84% of grammatical and 78% of ungrammatical stimuli. In the copy grammar, participants responded correctly to 82% of grammatical and 81% of ungrammatical stimuli. Results are reported in Table 2.11

Accuracy on grammatical stimuli Accuracy on grammatical stimuli by group and grammar is reported in Figure 2.13 (see also Tables 2.11 and 2.12). It is interesting to note how the performance on the finite state grammar was markedly different from the performance in the two supra regular grammars for the hearing group only.

Table 2.12: Hearing participants: group results

Grammar	Type of string	Mean accuracy (SD)
ab^na	Grammatical (N=2,3)	98 (10)
	Grammatical extension (N=4,6)	90 (29)
	Foil missing tile	99 (10)
	Foil wrong tile	90 (29)
ww^r	Grammatical (N=2,3)	91 (29)
	Grammatical extension (N=4,6)	74 (44)
	Foil missing tile	83 (37)
	Foil wrong tile	74 (44)
ww	Grammatical (N=2,3)	92 (27)
	Grammatical extension (N=4,6)	69 (46)
	Foil missing tile	86 (35)
	Foil wrong tile	77 (42)

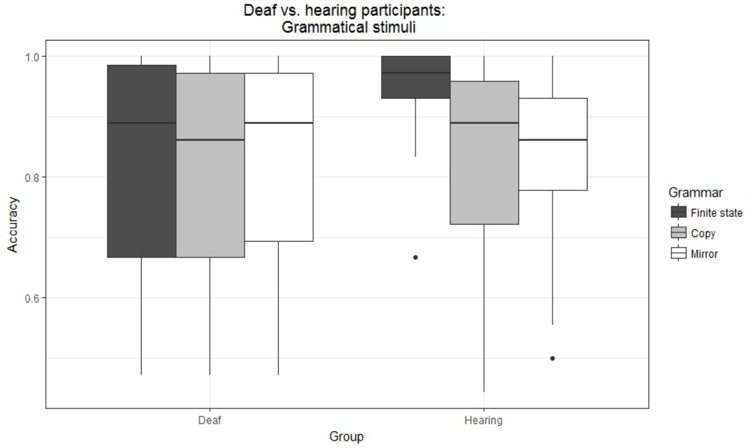


Figure 2.13: Six box plots representing accuracy on grammatical stimuli in the three different grammars, separated for Deaf (D) and Hearing (H) participants.

Accuracy was analyzed using generalized linear mixed models for binomially distributed outcomes with subjects and items as random intercepts. We entered two categorical predictors, Group (Deaf vs. hearing) and grammar (mirror grammar (reference level) vs. finite state grammar and copy grammar). The full model is represented in Table 2.13. The analysis confirmed the observation that the difference between Deaf and hearing participants was limited to the finite state grammar.

Table 2.13: Deaf vs. hearing participants - Grammatical stimuli: logistic regression analysis.

Fixed effects	β	SE	z value	
grammar ab^na	-0.08	0.18	-0.464	
grammar (ww)	-0.07	0.18	-0.37	
group	0.15	0.54	0.27	
grammar ab^n vs. group	1.66	0.30	5.445	***
grammar ww vs. group	-0.09	0.26	-0.36	

*** p<.001
Reference levels: grammar = mirror grammar;
group = deaf;

Accuracy on ungrammatical stimuli Accuracy on ungrammatical stimuli by group and grammar is reported in Figure 2.14 (see also Tables 2.11 and 2.12).

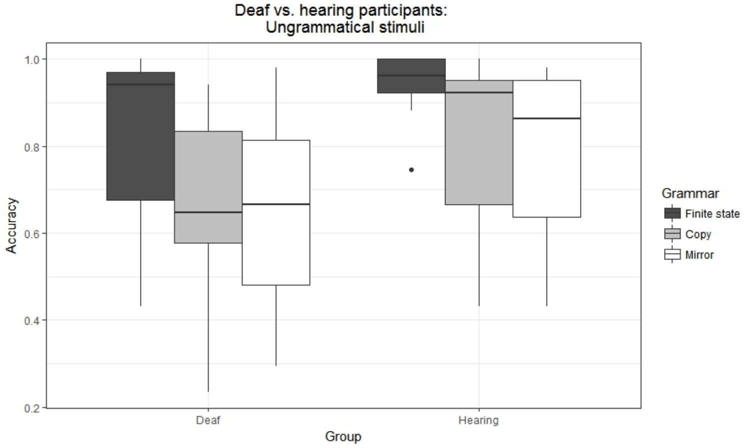


Figure 2.14: Six box plots representing accuracy on ungrammatical stimuli in the three different grammars, separated for Deaf (D) and Hearing (H) participants.

Accuracy was analyzed using generalized linear mixed models for binomially distributed outcomes with subjects and items as random intercepts. We entered two categorical predictors, Group (Deaf vs. hearing) and grammar (mirror grammar (reference level) vs. finite state grammar and copy grammar). The grammar (finite state grammar) by group interaction was significant. Accuracy on the finite state grammar was higher than in the mirror grammar, especially for the hearing group (as observable in Figure 2.14). Overall, accuracy was higher for the hearing than the Deaf people.

Individual above-chance performance Performances of participants of the two groups were analyzed individually, considering success in each grammar separately for type condition (Table 2.15): grammatical: at least 15/20 trials correct (exact binomial test p=.02); grammatical extension at

Table 2.14: Deaf vs. hearing participants - Ungrammatical stimuli: logistic regression analysis.

Fixed effects	β	SE	z value	
grammar ab^na	1.27	.14	9.010	***
grammar ww	0.07	0.12	0.567	
group	0.93	0.45	2.08	*
grammar ab^n vs. group	0.66	0.24	2.715	**
grammar ww vs. group	0.18	0.19	0.921	

* p<.05; ** p<.01; *** p<.001
Reference levels: grammar = mirror grammar;
group = deaf;

least 12/16 trials correct (exact binomial test p=.04); foil missing tile: at least 18/26 trials correct (exact binomial test p=.04); foil wrong tile: at least 18/25 trials correct (exact binomial test p=.02).

Table 2.15: Deaf vs. hearing participants - individual performance. Cells report the number of participants who performed above chance.

(15 participants per group)		Deaf participants			Hearing participants		
Condition	*Grammar*	ab^na	ww^r	ww	ab^na	ww^r	ww
Grammatical		13	13	14	15	13	14
Gramm ext		8	10	8	14	9	9
Foil missing		13	11	13	15	13	13
Foil tile		11	6	6	14	9	11

Almost all participants could recognize the patterns presented during familiarization (Grammatical stimuli vs. foil missing). Almost all hearing participants showed the acquisition of the finite state grammar, being able to accept grammatical extensions. The same is true for only half of the Deaf participants. About half participants acquired the mirror and the copy grammar, and this is true for both Deaf and hearing participants.

With the Fisher's exact test for count data we compared, for each grammar and type, if the number of above-chance participants differed between the Deaf and the hearing group: the only significant difference was in grammatical extensions in the finite state grammar (p=.035).

Correlation between accuracy and the Corsi score In order to investigate the role of memory in the VAGL task, we performed a correlation analysis between accuracy on grammatical and ungrammatical stimuli in the three grammars, separated for Deaf and hearing participants. Correlation coefficients and p values are reported in Table 2.16.

Table 2.16: Deaf vs. hearing participants - correlation between accuracy and the Corsi score for grammatical (G) and ungrammatical (U) stimuli

		Deaf participants		Hearing participants	
	Grammar	r	p	r	p
	ab^na	-.46	.08	.23	.41
G	ww^r	.05	.86	.54	.04
	ww	-.19	.51	.64	.01
	ab^na	.01	.97	.41	.13
U	ww^r	.30	.28	.59	.02
	ww	.02	.95	.47	.08

Interestingly, visuo-spatial span seems to have a big role for hearing participants considering the two supra-regular grammars, but this is not the case for the Deaf participants. On the contrary, in the Deaf group the visual-spatial span was inversely correlated to accuracy on grammatical stimuli in the finite state grammar, but the correlation was not significant.

Discussion

Deaf and hearing participants: different strategies? The question if Deaf and hearing participants used different strategies to solve the task arises from one observation: The Corsi score correlated with accuracy in the supra-regular grammars considering the hearing participants (with the exception of accuracy on ungrammatical stimuli in the copy grammar, even though r is quite high, .47, and $p<.1$) but not considering the Deaf participants. This fact suggests therefore that the two groups might have performed the task following different strategies. Future research should focus on the mechanisms that are at play during the familiarization phase, and maybe it might be useful to ask directly the participants for the rule that they think can produce the familiarization sequences, instead of asking for grammaticality judgments. Or, future studies might ask participants to produce strings of a given grammar starting from the vocabulary. Moreover, brain-imaging techniques should provide evidence on the neural mechanisms involved in the VAGL task, to see to what extent there is overlapping between hearing and Deaf participants and to what extent there is not.

A second important observation is that, considering supra-regular grammars, even if the number of above-chance participants did not differ between groups, accuracy was overall lower in the ungrammatical stimuli in the Deaf group. This might reflect differences linked to attention (sustained attention is needed to perform the task. Sequences followed each other without breaks and even one second of distraction would have lead to possible errors).

A further question that arises from the present data is why about 50% of Deaf participants could not generalize to correct stimuli of N=4 and N=6 in the finite state grammar. This, plus the observation that the correlation coefficient between the Corsi score and accuracy on ab^na grammatical stimuli was negative, in the Deaf participants (Table 2.15. Even if the correlation was not significant, the correlation coefficient was quite high) might suggest that inside the Deaf group participants used different strategies to process the finite state grammar. Specifically, deaf participants with higher Corsi might have counted the number of *b* elements between the two *a*s, acquiring so a subset of the rule, whereas participants with lower Corsi might have inferred a rule indicating that the *b*s between the two *a*s should have been 2, 3 or *"more than 3"* (from N=5). But this is only speculation, and as I have already stated, future AGL studies should inquire about the acquired rule and the strategies used to perform the task.

Limitations I acknowledge as a limit of the present experiment the fact that not all the Deaf participants were native signers. Nevertheless, i) non-native signers were exposed to LIS during childhood, ii) they were all proficient signers, i.e. they all had learned a full linguistic system (LIS), iii) they were all active members of some Deaf institute, using therefore LIS as everyday mean of communication. Those points are important: A recent study demonstrated that late sign language exposure affected the ability to perform syntactic judgments on American Sign Language (ASL) in children and adolescents. Nevertheless, non-native signers with adequate exposure to ASL performed as native signers and in non-native signers as in native signers performance improved with age (Novogrodsky et al., 2017).

2.4.8 General discussion

Computational vs. cognitive complexity If cognitive complexity and computational complexity overlapped, the mirror grammar would be easier to process than the copy grammar. This was not the case: Our results tease human computations and the Chomsky hierarchy apart. In this respect, one might draw a parallelism from Tversky and Kahneman work on reasoning fallacies. Probability theory states it clear that the probability of a conjunction, P(A∧B), is never higher than the probability of its constituents, P(A), P(B). Nevertheless, when specific scenarios are given, humans happen to judge P(A∧B) more likely than P(A) (see e.g. the famous Linda's problem, Tversky and Kahneman 1983). Humans use intuitive heuristics that are not subject to probability theory. In the same vein, it might be possible that cognitive mechanisms involved in AGL are guided by heuristics that totally diverge from that of formal language theory. If this was the case, linguistics and cognitive science might not benefit much from the work stemming from Chomsky (1956). On the contrary, what I think one should ask is why the Chomsky hierarchy is not a direct measure of cognitive complexity, and what factors do play a role during grammar processing that move human cognition away from what is predicted by the theory of computation. As a matter of fact, this approach guided the present study, and we were not expecting the mirror grammar to be easier than the copy grammar, but the opposite.

Working memory? As already described, previous investigation with natural languages highlighted a human preference for crossed over nested dependecies (Bach et al., 1986). In light of those results, we were actually expecting the copy grammar to be easier than the mirror grammar. On the contrary, we observed no significant differences between the two supra-regular grammars. Two different explanations could account for the present data. The first explanation takes into account working memory. I will now focus on this first explanation and I will examine the second one in the next section.

As presented in the introduction of the present study, Bach et al. (1986) results have been explained thanks to the introduction of embedded pushdown automata (Joshi, 1990), according to which the working memory load needed to compute nested dependencies is higher than that required to compute crossed dependencies. In the human working memory limitations should therefore reside one of the explanations about why cognitive complexity and computational complexity do not overlap. Therefore, a possible justification for our results could be the reduced working memory load that characterized the present paradigm. The lower working memory engagement, compared to (Bach et al., 1986) might indeed have neutralized the copy grammar advantage.

At this point, one might ask how cognitive processing and computational processing could overlap. This might be possible to minimize the load of working memory through the manipulation of the experimental setting AND giving human more memory resources. In this way, it might be possible to observe AGL results that follow what predicted by the Chomsky hierarchy. This prediction might be tested either with transcranial direct current stimulation (tDCS) or repetitive transcranial magnetic stimulation (rTMS). To date, both techniques have been used to enhance performance in AGL studies (De Vries et al. 2010 for tDCS and Uddén et al. 2008 for rTMS), but with regular Reber-like grammars only. I argued that the comprehension of simple and complex syntactic structures might be improved through anodal tDCS in a recent work with Alessandra Vergallito, Carlo Cecchetto, Erica Varoli, and Leonor J. Romero Lauro (Giustolisi et al., 2018).

A role of visual symmetry? The second possible explanation involves a geometrical concept. Presenting the mirror grammar in the visual modality resulted in dealing with a peculiar visual process: symmetry. In fact, the mirror grammar generated mirror symmetric patterns, in which the second half reflects the first half. This might be an interfering factor if participants processed strings as a whole, and not sequentially, element by element. It is well known, in fact, that mirror symmetric patterns are detected easier than repetition patterns (or non-reflected identity, as in the copy grammar. See Bruce and Morgan 1975). Mirror symmetry perception is indeed "*effortless, rapid and spontaneous*" (Wagemans, 1997). Moreover, we are used to mirror symmetry, being it pervasive in the natural and in the artificial word (Treder, 2010) [2]. If symmetry played a role in the present study, this might indeed explain the different results between the present study and previous studies highlighting a preference of crossed dependencies over nested dependencies (Bach et al., 1986). It might actually be the case that humans favor the former, but the special symmetric status of mirror stimuli made them easier to be processed than expected if no symmetry was involved, therefore darkening other effects linked to the type of dependecies. The presentation modality we used, though, should have avoided this risk: Firstly, the stimuli were presented sequentially. Secondly, the whole pattern appeared on the screen just for a brief moment. So, even if I acknowledge that visual symmetry might have had a role in the present task and I cannot rule out this possibility with the present paradigm, I do not favor this rather cumbersome explanation.

The Deaf results: VAGL and the ASH, limits and future directions From the present data, two major conclusions can be drawn. Firstly, Deaf signers can processes complex visual sequences and recognize them (as shown by the good performance with N=2 and N=3 in all grammars, see Table 2.15, but also Table 2.8). Moreover, with the appropriate experimental conditions (experiment 4) some of them (about 50%) can extract rules at the context-free and mildly-context-free level of the Chomsky hierarchy with only a couple of minutes of exposure to the target grammar. Overall, this argues against the ASH, showing that Deaf people can learn complex algebraic patterns and nested tree structures. The unexpected results of a different role of visuo-spatial span between Deaf participants

---

[2] Consider, e.g. humans and other animals, as well as objects like a pair of glasses...

and hearing participants suggests that Deaf people might perform the task using different strategies than hearing participants, but further research should examine in depth this issue.

Secondly, everyday use of a sign language instead of a spoken language does not augment people's ability to acquire an artificial language presented in the visual modality. This argument has a limit, though. Even if the hearing participants were not signers, and had no exposure to any sign language, they also had extensive experience with a visual language, i.e. the written form of their spoken language (Italian, in this case) [3]. Future research should explore this issue focusing on the acquisition of artificial languages by Deaf signers using sign-language linguistic stimuli, e.g. pseudo-signs.

---

[3] Nevertheless, the acquisition of LIS by the Deaf signers occurs in a natural way, i.e. implicitly, whereas literacy is taught explicitly.

CHAPTER 3

Visual statistical learning

In this chapter I leave behind algebraic patterns and nested tree structures to focus on chunking and ordinal knowledge (see Figure 1.3). Specifically, I investigated the ability to group together continuous items that recur as a whole and to recognize what item occurs in what order. I did so with a visual statistical learning experiment. This experiment was designed and data were collected from March to September 2017 at the San Diego State University, Laboratory for Language and Cognitive Neuroscience directed by Prof. Karen Emmorey. Ms. Cindy O'Grady Farnady helped me in recruiting and testing Deaf participants.

3.1 Statistical learning

Statistical learning involves a set of mechanisms that work in different modalities and through which we can encode regularities across space and time. Those mechanisms operate implicitly (Frost et al., 2015). As for artificial grammar learning (AGL, see chapter 2.1), studies assessing implicit statistical learning are typically composed of two phases: familiarization and testing. Without receiving any explicit information, participants are exposed to some kind of stimulus regularity (familiarization phase). Then, learning the (unmentioned) regularity is assessed (testing phase). If statistical learning has occurred, participants should be able to discriminate between familiar and non-familiar stimuli (e.g. familiar vs. non-familiar triplets of stimuli, see Arciuli and Simpson, 2012; Saffran et al., 1996a), and they should be faster/more accurate to perform actions on familiar compared to non-familiar stimuli (e.g. serial reaction time paradigm, Nissen and Bullemer, 1987).

The term "statistical learning" has been in use for over 20 years, starting from the seminal paper by Saffran et al.*"Statistical learning by 8-Month-Old Infants"*, published in Science in December 1996. The rationale behind that statistical learning study was to show that language acquisition is constrained not only by innate mechanisms (see e.g. Chomsky, 1980; Piattelli-Palmarini, 1994), but also by experience-dependent learning mechanisms. Specifically, Saffran et al. (1996a) focused on those mechanisms that allow infants to segment words from the continuous speech stream. They proposed that we, humans, are very good in detecting statistical regularities within and between words and that this ability should help infants in the acquisition of the lexicon (see also Saffran et al., 1996b). As an over-simplification, if we consider the phrases *simple garden* and *simple jacket*, the transitional probability from one syllable to another within a word is greater than the transitional probability from the last syllable of the adjective and the first syllable of the following noun (i.e. given *sim* there is always *ple*, but after *ple* there can be either *gar* or *ja*). In a first experiment, Saffran et al. (1996a) exposed 8-months old American infants to a continuous speech stream. The continuous speech stream was composed by 3 invented words, each made by 3 syllables. The transitional probabilities between syllables varied: 1.0 between words and 0.33 within word. This was the only cue to word boundaries. After a short familiarization phase (2 minutes) participants were able to distinguish words from non-words showing greater fixation times (corresponding to listening times) for non-words than for words. Non-words were composed by the same syllables as words, but combined in a way that was new for the infants. In a second experiment, non-words corresponded to strings that were heard during the familiarization phase (between words transition). Using the *"simple garden"* example, in experiment 1 *"simgar"* but not *"plegar"* was a non-word, whereas in experiment 2 non-words were of the *"plegar"* type. Also in experiment 2 participants could distinguish words from non-words.

Building on this first study, subsequent experiments have confirmed that from infancy to adulthood humans are sensitive to the co-occurrence of items not only with verbal, but also with non-verbal stimuli (e.g. Gebhart et al., 2009) and that this sensitivity persists across modalities (e.g. visual: Bulf et al., 2011).

3.1.1 Statistical learning and grammar learning

As I have already explained, statistical learning and artificial grammar learning experiments are very similar: They both involve the distinct phases of familiarization and testing. However grammar learning, as opposed to statistical learning, involves abstraction of patterns that can be generalized to elements that have not been presented during familiarization. Considering a Reber's like grammar (Reber 1967, see chapter 2.1), statistical learning and artificial grammar learning are two sides of the same coin (see Perruchet and Pacton, 2006). In fact, to perform an AGL experiment with a Reber's like grammar participants need to learn the transitions between stimuli (see Figure 2.1) and they might be able to do so without extracting any kind of abstract rule. What participants store in memory to perform these tasks are sequences of items. On the contrary, the crucial difference between statistical learning and grammar learning is that in the latter what participants should store in memory are abstract schemata. This is represented in Dehaene et al.'s taxonomy (2015), with algebraic patterns and nested tree structures on one side and the lower levels on the other side.

Considering the statistical learning / grammar learning literature, and going again back in time to the foundation studies in the field, the fact that infants are able to extract algebraic rules without relying on statistics properties of the stimuli was demonstrated by Marcus et al.. Their paper, *"Rule Learning by Seven-Month-Old Infants"*, was published in Science in January 1999. During the familiarization phase, participants (7 months-old infants, 1 month younger than in Saffran et al.'s 1996a experiment) were familiarized for 2 minutes (as in Saffran et al., 1996a) to one of two abstract rules: ABA or ABB. The ABA grammar generated strings of three syllables, with the first and the third syllable alike. The ABB grammar generated strings of three syllables, with the second and the third syllable alike. The words presented during the familiarization phase were different from those presented in the testing phase. Infants could distinguish words from non-words regardless of phonetic similarities between A and B elements. Moreover, they could distinguish between ABB words and AAB non-words, showing that they could acquire the whole grammar structure, and not only a subset (e.g. element repetition). Marcus et al.'s study (1999) had the merit to highlight that statistical learning mechanisms cannot explain alone the process of language acquisition. In fact, after Saffran et al. (1996a), some commentaries tried to reduce language acquisition to mere statistical computations (e.g. Bates and Elman, 1996).

3.1.2 The flourishing of statistical learning studies

In recent decades, studies on statistical learning have grown exponentially, which is justified by the recognized role of statistical as a learning mechanism involved in almost every cognitive process (Perruchet and Pacton, 2006), but particularly for language. A growing body of evidence suggests that some statistical learning mechanism supports literacy development. Such statistical learning mechanisms are argued to sustain the recognition of the probabilistic patterns of association within orthographic representations and between graphemes and phonemes, building a fundamental scaffolding for developing of reading and spelling skills (Treiman and Kessler, 2006).

Different studies reported a relationship between statistical learning and reading ability in both children and adults (Arciuli and Simpson, 2012; Spencer et al., 2015) and in both first (L1) and second (L2) languages (Frost et al., 2013). For example, Arciuli and Simpson (2012) showed a positive correlation between statistical learning and L1 reading proficiency as measured by the reading subtest of the Wide Range Achievement Test 4 (Wilkinson and Robertson, 2006), which investigates the ability to read aloud different orthographic strings. Frost et al. (2013) found that, in English L1 speakers, statistical learning positively correlated with the ability to learn the words structural properties of Hebrew (L2), a Semitic language. Crucially, the structure of Hebrew words (spoken and printed) differs from that of English words. As all semitic languages, Hebrew words present a 3-consonants root expressing the core meaning, whereas related nouns, adjectives, verbs etc. are composed adding abstract phonological configurations of vowels or vowels and consonants. Therefore, written Hebrew words consist basically on consonants, with vowels superimposed as diacritics.

In addition to these studies, the link between statistical learning and literacy is supported by the finding that individuals with dyslexia seem to show statistical learning impairments (e.g. Gabay et al., 2015; Sigurdardottir et al., 2017; but see Rüsseler et al., 2006).

3.2 Study 2: statistical learning in deaf and hearing adults

In the present study, we focused on the link between statistical learning and literacy considering a population that, to the best of our knowledge, has never been considered before in this regard: congenitally Deaf adult signers. First, we were guided by the following consideration: computational models taking a statistical learning approach to literacy have been developed to represent the behavior of hearing individuals, and studies linking statistical learning abilities and written language proficiency have only involved hearing readers. With this in mind, one might wonder if the same statistical learning – reading/writing relation occurs in Deaf people. This question is interesting for several reasons. In hearing people, orthographic processes assume a mapping between sound-based phonological representations and orthographic representations. From a developmental perspective, this mapping develops from preexisting phonological representations of spoken language to not-yet-known orthographic representations. Considering the case of Deaf readers/writers means considering the case of people who have no or partial phonological representations of speech that must be mapped onto orthographic words (Goldin-Meadow and Mayberry, 2001). We hypothesized that variation in statistical learning could partially account for the high variability in reading proficiency in the Deaf population (Qi and Mitchell, 2011), and our primary goal was to test this hypothesis. To do so, we ran a visual statistical learning experiment, and we collected several measures assessing reading [1], spelling, American Sign Language (ASL), and cognitive skills in order to perform a correlational analysis. In particular, we investigated whether a possible association between statistical

---

[1] Assessing reading aloud abilities is not appropriate for Deaf individuals, so we used a reading comprehension task.

learning and reading ability might be mediated by a more general relationship between statistical learning and natural language ability (Arciuli and Simpson, 2012), in this case ASL skill.

Moreover, studying statistical learning in deaf adults is of great interest in light of the debate concerning the auditory scaffolding hypothesis (ASH, Conway et al., 2009; Hall et al., 2017; von Koss Torkildsen et al., 2018). As I explained in chapter 1, the ASH states that learning and producing sequential information might be impaired in deaf individuals, because the development of those abilities are sustained by hearing experience. Evidence in favor of this hypothesis comes from studies demonstrating sequence learning/processing deficits in deaf children with cochlear implant (e.g. Conway et al., 2011b). However, the ASH has been challenged by a study in which deaf adults outperformed hearing adults in a visual rhythmic task, which was highly sequential in nature (Iversen et al., 2015). In addition, Hall et al. (2017) failed to replicate Conway et al. (2011b). Specifically, Hall et al. (2017) found no evidence of sequence learning in either deaf or hearing children using Conway et al.'s task, and they found evidence of similar sequence learning in both groups using a different task (a serial reaction time paradigm). Recently, von Koss Torkildsen et al. (2018) also reported similar visual statistical learning in deaf and hearing children (aged 7 – 12 years). Moreover, even if with great variability, as I showed in chapter 2, Deaf adults can perform an artificial grammar learning task with stimuli presented sequentially.

The further goal of the present study was to test the ASH for congenitally Deaf adults using a statistical learning experiment, therefore analyzing a different level of sequence encoding than in chapter 2. In light of this hypothesis, congenitally Deaf adults should perform worse than hearing adults in a sequential statistical learning task because of their lifelong lack of hearing experience.

Methods

Participants Twenty-five Deaf ASL signers and 27 hearing non-signers (native English speakers) participated. Through a background survey, all participants reported no history of language impairment. Deaf participants were all native signers (born into deaf signing families) or early signers (ASL was acquired before age 6); they all used ASL as their primary means of communication and written English as alternative means of communication. They were all born deaf, with either severe (71–90 dB) or profound (90–120 dB) hearing loss. One deaf participant and two hearing participants were excluded from the analysis because of lack of attention during the statistical learning familiarization phase, whereas one hearing participant was excluded because he was unwilling to complete the spelling assessment. The final sample consisted of 24 Deaf participants (mean age=32.5, SD=8.3; mean years of education = 16.5, SD= 3.0; 13 females) and 24 hearing participants (mean age= 30.9, SD=13.2, mean year of education = 15.6, SD= 1.9; 13 females). The two groups did not differ significantly on age (t(38.77)=0.48, p=.63), and level of education (t(38.80)=1.44, p=.16).

The experiment took place at San Diego State University. All Deaf participants received a monetary reimbursement for their participation. Hearing participants received either course credit or a monetary reimbursement for their participation. The Institutional Review Board of San Diego State University approved this study.

All participants underwent an assessment battery that measured print exposure, English reading and spelling skills, nonverbal IQ, and ASL skills (deaf participants only). The battery included the following tests:

Author Recognition Test (ART; Acheson et al., 2008). This test measures print exposure by asking participants to recognize names of authors presented in written form. Scores are computed as the number of hits (correctly identified authors) minus false alarms (incorrect identifications).

Maximum score is 65.

Peabody Individual Achievement Test Revised (PIAT-R; Markwardt 1989) – Reading comprehension subtest. This task measures reading comprehension by asking participants to silently read a sentence and choose among four pictures the one that best matches the sentence. While performing the decision, the sentence is not visible. In this task, vocabulary level progressively increases. Maximum score is 100.

Spelling recognition test (S-rec, Andrews and Hersch 2010). This test measures spelling skills by asking participants to identify incorrectly spelled words from a list of 88 words (half correctly spelled and half misspelled). The test score is calculated as the number of correctly classified items, both hits and correct rejections. Maximum score is 88.

Spelling Production task (S-pro). This task measures spelling abilities by asking participants to type words using a Cloze procedure in which a sentence context is provided and the first letter of the target word is presented (e.g., *In the US, temperature is measured in degrees F....*). Maximum score is 30.

Kaufman Brief Intelligence Test – Matrices (KBIT-2; Kaufman and Kaufman 2004). This task assesses non-verbal intelligence through a visual pattern completion task. Maximum score is 46.

ASL Comprehension Task (ASL-CT; Hauser et al. 2015). This task assesses ASL comprehension skills through a 30-item multiple-choice task (matching between a signed stimulus and one of four drawings/videos or a drawing/video and one of four signed stimuli).

ASL Sentence Repetition Task (ASL-SRT; Supalla et al. 2014). This test assesses ASL fluency by asking participants to repeat pre-recorded ASL sentences of increasing complexity. The maximum score is 35.

Materials

Stimuli Stimuli were 16 visual shapes taken from Fiser and Aslin (2001). They were organized into eight triplets following Siegelman et al. (2017). Specifically, four triplets, made from four shapes, had between-shapes transitional probabilities of .33, whereas four triplets, made from the remaining twelve shapes, had between-shapes transitional probabilities of 1. Labelling the shapes with a number from 1 to 16, four triplets were made from four shapes (e.g. 1-2-3; 2-1-4; 4-3-1; 3-4-2) and four triplets were made from 12 shapes (e.g 5-6-7; 8-9-10; 11-12-13; 14-15-16).

Familiarization During the familiarization phase, eight triplets appeared on the screen in a pseudo-random order (the same triplet never appeared twice in a row). Each triplet was repeated 30 times, for a total duration of about 10 minutes. Each shape appeared on the screen for 400 ms, with an inter-stimulus interval of 250 ms. Within triplets, the shapes appeared at three different screen locations, as shown in Figure 3.1. This presentation modality is similar to the transient modality of Study 1.

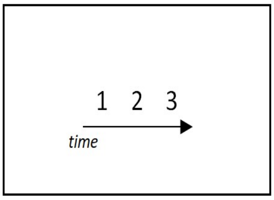


Figure 3.1: Presentation modality

1, 2, 3 correspond to the sequential location of different shapes. The three shapes making a triplet were never visible all together on the screen: each shape appeared on the screen for 400 ms, and the following shape appeared 250 ms later.

Testing The testing phase followed Siegelman et al. (2017). In short, two sub-scores composed the final score: pattern recognition score and pattern completion score. Pattern recognition was composed of 34 trials, randomly presented to each participant. In each trial, participants had to select the familiar stimuli among a list of four or two. Pattern completion was composed of eight trials randomly presented to each participant. In each trial, participants had to select among three shapes which one completed an incomplete pattern. The maximum visual statistical learning score was 42 (34 pattern recognition + 8 pattern completion). Individual above chance level was success on 23 or more trials (Siegelman et al., 2017).

Overall procedure Deaf participants completed the visual statistical learning task and the assessment tests in separate sessions on different days. Hearing participants completed the visual statistical learning Task and the assessment tests in one session lasting about 60 minutes. The visual statistical learning task was always administered first.

Results and analysis

Table 3.1 reports assessment raw scores. Crucially, Deaf and hearing participants did not differ significantly on nonverbal IQ. Spelling recognition scores did not differ significantly between groups, but the hearing group outperformed the Deaf group on the reading comprehension task (PIAT-R), on the Author Recognition Test (ART), and on the spelling production task. Performance on the ART was highly correlated with age in the hearing group (r=.60, p=.002), suggesting that this test might not be an appropriate measure of print exposure in adult participants who vary in age. Therefore, we considered this measure for the Deaf group only.

The mean visual statistical learning score for the Deaf participants was 26.5/42 correct (SD=5.3) and for the hearing participants it was 24.5/42 correct (SD=4.7). This difference was not significant (t(46)=1.38, p=.17, Cohen's d=.40). Individual scores are reported in Figure 3.2. Individual chance level was set at 23 correct trials, following the criterion proposed

Table 3.1: Means and SDs for assessment scores for Deaf and hearing participants

N=48	Deaf M (SD)	Hearing M (SD)	t	
Print exposure (ART)	10.9 (7.8)	16.7 (11.8)	-2.04	*
Reading comprehension (PIAT-R)	79.3 (12.3)	91.0 (4.7)	-4.36	***
Spelling recognition (S-rec)	72.9 (6.8)	74.8 (6.8)	-0.97	
Spelling production(S-pro)	0.70 (0.20)	0.82 (0.12)	-2.61	*
Nonverbal IQ (KBIT)	105.7 (11.4)	108.8 (13.0)	-0.88	
ASL comprehension (ASL-CT)	0.88 (0.09)	-	-	
ASL repetition (ASL-SRT)	22.7 (5.0)	-	-	

* $p<.05$; ***$p<.001$

by Siegelman et al. (2017) (obtained through a computer simulation). In both groups, the majority of participants performed above chance (Deaf participants: 18/24; hearing participants 15/24).

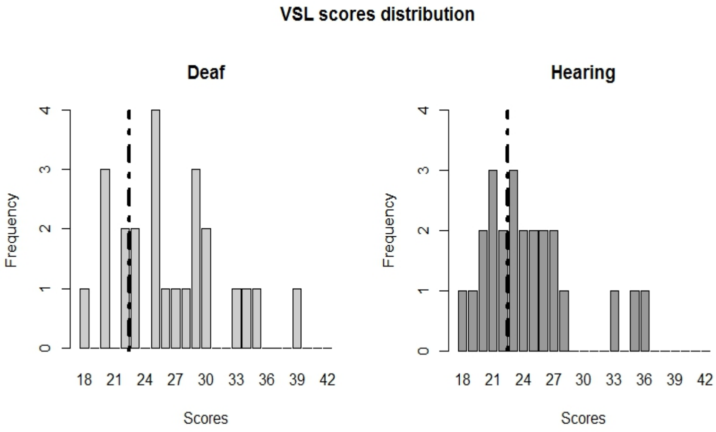


Figure 3.2: Visual statistical learning scores distribution in deaf (left) and hearing (right) participants.
The dashed black line indicates chance level (<23).

To examine the relationship between visual statistical learning scores, demographic characteristics (age and education), and assessment scores we performed a correlation analysis. Table 3.2 shows correlation coefficients (r) and p-values. The correlation between visual statistical learning scores and demographic information was not significant for either group. Similarly, for both groups there were no significant correlations between visual statistical learning scores and spelling scores. The correlation between visual statistical learning scores and KBIT scores was positive in both group, but not significant. visual statistical learning scores positively correlated with reading comprehension scores in the hearing group (r=.44; p=.03) and in the Deaf group, although this correlation was not significant (r=.30; p=.16). In the Deaf group, visual statistical learning scores positively correlated with ART scores, but the correlation was not significant (r=.37, p=.07). Finally, for the Deaf participants, the correlation between visual statistical learning scores and ASL proficiency (ASL-SRT or ASL- CT scores) was not significant.

To further examine the relationship between visual statistical learning scores and reading comprehension scores in the two groups, we performed a linear regression analysis with the visual statistical learning score as dependent variable (see Figure 3.3). As independent variables, we considered group (deaf vs. hearing), the PIAT-R score, and their interaction. All predictors were entered in the model mean centered. The interaction was not significant, and it was removed without decreasing the model's goodness of fit (F=1.93, p=.17). The main effect of PIAT-R was significant (β=0.16, SE=0.07, t=2.33, p=.02) and the main effect of group was significant (β=4.44, SE=1.54, t=2.87, p=.01). This pattern indicates that when PIAT-R scores are taken into account, accuracy in the visual statistical learning task is higher for the Deaf group than the hearing group.

Table 3.2: Correlations between visual statistical learning scores and biographic characteristics and assessment scores in the two groups

		Age	Edu	PIAT-R	S-rec	S-pro	KBIT	ART	ASL-CT	ASL-SRT
D	r	-.23	-.11	.30	.12	.-04	.20	.37	.13	.13
	p	.28	.62	.07	.16	.44	.34	.07	.55	.54
H	r	-.09	-.16	.44	-.02	.04	.37	-	-	-
	p	.68	.45	.03	.92	.84	.08	-	-	-

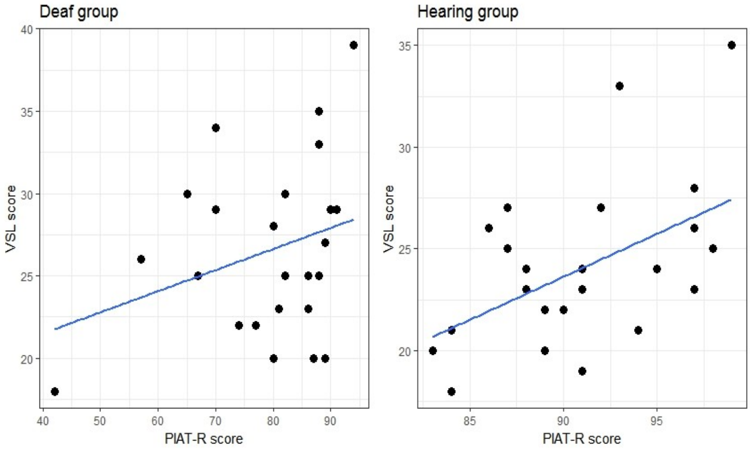


Figure 3.3: Visual statistical learning score and PIAT-R score scatter plot
Scatter plot with best-fitting regression line showing visual statistical learning scores as a function of PIAT-R scores in the deaf (left) and in the hearing (right) groups.

Discussion

In this study, we used a visual statistical learning task with stimuli presented sequentially across time and space to assess statistical learning skills in deaf and hearing adults. The first goal was to examine whether statistical learning abilities correlated with reading ability in deaf adults. Visual statistical learning scores positively correlated with reading comprehension (PIAT-R) scores in the hearing group (r=.44; p=.03). The correlation was also positive in the Deaf group (r=.30), but it was not significant. Overall, our results are consistent with those of Arciuli and Simpson (2012) who found a positive correlation between statistical learning and oral reading skill in hearing children and adults. Measuring the ability to read out loud is not appropriate for deaf individuals who do not use speech, and we therefore used a reading comprehension task (PIAT-R) to assess reading skill. Performance on this task is tightly linked to vocabulary knowledge, as the level of vocabulary difficulty increases trial by trial. Therefore, the present results with hearing adults are also in line with previous research reporting higher statistical learning abilities in children with greater vocabulary knowledge (Evans et al., 2009).

An interesting finding is that, despite the fact that our group of Deaf participants had lower reading skills than the hearing participants, the two groups did not differ on visual statistical learning scores. As indicated by the linear regression analysis, when reading comprehension scores are taken into account, visual statistical learning scores are higher for the deaf than the hearing group. To our opinion, this might indicate that only Deaf people with very high statistical learning abilities might become highly proficient readers. This observation is of great interest and needs further investigation, especially in light of the possibility that training statistical learning skills might have an impact on reading skills. Of course, this is just a speculation and further factors might play a role. Only longitudinal studies assessing statistical learning abilities before and after literacy instruction could prove this line of reasoning.

As an additional note, we suspect that the absence of a correlation between ASL skill and visual statistical learning scores in the Deaf group might be a consequence of the sequential presentation modality of the present paradigm. Previous research has indicated that the strategies for segmenting a signed language differ from those used to segment spoken languages (Brentari, 2006). We hypothesize that scores on a statistical learning task with stimuli presented simultaneously instead of sequentially might correlate with ASL skill because the statistical learning mechanism involved in extracting simultaneous visual patterns might be particularly relevant for parsing sign language input. Further, it might be that statistical learning with stimuli presented simultaneously might correlate more strongly with reading skills in Deaf signers. Further research is needed to test these hypotheses.

As for the ASH (Conway et al., 2009), using a sequential visual statistical learning task, we found no significant difference between the performance of Deaf and hearing participants. The prediction of the ASH is that Deaf individuals should perform worse than hearing individuals in tasks tapping temporal and sequential order, and this prediction was not met. According to our interpretation, the present result suggests that there is no need for auditory scaffolding to develop sequencing skills. However, another possibility is that deaf adults have had enough time throughout their lives to develop a set of strategies based on declarative memory to compensate for early deficits in temporal sequencing skills, as has been proposed for individuals with dyslexia (Lum et al., 2014). With regard to the Deaf population, this second hypothesis is weakened by the results of Hall et al. (2017) showing comparable visual statistical learning skills between Deaf children who were native signers (and thus were not language-deprived) and hearing children.

In summary, the results of our study indicated that visual statistical learning skills with stimuli presented across space and time are comparable between Deaf and hearing individuals matched in age, education, and IQ. Moreover, we showed that when reading scores are taking into account, Deaf participants perform better than hearing participants.

CHAPTER 4

Predictive timing

The main topic of the present chapter is the relationship between hearing experience and timing behavior. Differently from the previous chapters, the population under investigation was that of deaf children with cochlear implant (CI) with no exposure to sign language. The experimental study was designed with Prof. Maria Teresa Guasti and Prof. Natale Stucchi. Data collection was conducted at the Department of Otorhinolaryngology of the Hospital of Piacenza in collaboration with Dr. med. Domenico Cuda and Dr. Letizia Guerzoni.

4.1 Temporal information and predictive coding

The unconscious encoding of sequences on the basis of the temporal transitions between items is characterized by the brain emission of a mismatch response when an expected stimulus is missing or replaced by a different one (Dehaene et al., 2015). The mismatch negativity, described for the first time in the late 1970s (Näätänen et al., 1978), is a negative event-related potential component that peaks 100/200 ms after the onset of a sound that deviates from a regular auditory sequence (see Näätänen, 2003; Wacongne et al., 2012). Later on, analogous forms of this component have been found in other domains, e.g. visual (Pazo-Alvarez et al., 2003) and olfactive (Pause and Krauel, 2000). A validated interpretation of the mismatch negativity is that, when processing regular inputs, we build an abstract model that we can use to make predictions about forthcoming events. The mismatch signal should indicate the inconsistency between the prediction and the perceived event.

From the behavioral perspective, an evidence of this predictive coding consist in the human abilities to tap to a metronome. In fact, humans can build a model of the metronome cadence and use it to anticipate forthcoming events and tap in synchrony with them (e.g. Iversen et al., 2015; Zarco et al., 2009).

4.1.1 Timing processing and deafness

In Conway et al.'s view (2011a, but see also Conway et al. 2009, 2011b) the processing of temporal information is primarily sustained by hearing experience (auditory scaffolding hypothesis, ASH). That the lack of hearing experience might have an impact on the temporal domain was suggested also by further research (e.g. Bolognini et al., 2012; Kowalska and Szelag, 2006). For example, Bolognini and collaborators showed that deaf adults (congenitally deaf, 7 signers and 2 non-signers) performed worse than hearing participants in a task assessing the ability to discriminate the duration of a tactile stimulation on the index finger. Comparing the effect of transcranial magnetic stimulation on the primary sensorymotor area and on the auditory associative cortex, the Authors found that spatial versus temporal processing were sustained by different neuronal populations in the hearing group, but not in the deaf. This should reflect cross-modal functional reorganization of the auditory cortex. Intriguingly, these results are ad odd with Iversen et al. (2015), who found that Deaf adults synchronize to visual flashes better than hearing adults, as I will describe in the introduction of study 3. Iversen and coauthors discussed this issue suggesting a different impact of hearing experience on synchronization processes compared to other temporal processes. The focus of study 3 was timing skill involved in synchronization.

4.1.2 Motor sequencing impairments and language delays in deaf children with CI

The importance of processing temporal information is shared between speech production/comprehension and complex motor actions (Lashley, 1951; Zarco et al., 2009, and see also chapter 1). To study the effect of hearing deprivation on non verbal sequential abilities and timing skills, Conway et al. (2011a) assessed 24 prelingually profoundly deaf children with CI implanted before the age of 4 with a series of non verbal tasks. Four tasks were chosen from the NEPSY battery (Korkman et al., 1998):

Fingertip tapping to measure sensory-motor sequencing skill (participants were asked to tap the index finger against the thumb separately for each hand, and all the fingers sequentially against the thumb);

Finger discrimination to measure tactile perception (without visual feedback, participants were touched on one or two fingers and they had to recognize what finger was touched);

Knock and tap to measure manual response inhibition (participants learned different motor behaviors to use in response to specific examiner's behavior);

Design and copy to measure visual-motor integration and visual-spatial processing.

One task, the *dot location*, was a subtest of the Children's Memory Scale (Cohen, 1997) and it was used to measure non-verbal visual-spatial memory by asking children to reproduce several dots patterns.

The performance of the group of children with CI was compared to that of a group of 31 hearing peers with better forward and backward digit span and higher receptive vocabulary scores. There was no between group difference on the finger discrimination task, the knock and tap task and the dot location task. The hearing group performed better than the CI group in the design and copy task, but the hearing group performed better than expected considering normative data, whereas children with CI were age-appropriate.

As reported by the Authors, children with CI performed worse than hearing children in the fingertip tapping task. To be more precise, children with CI were slower than hearing children in tapping their index finger against the thumb with their non-dominant hand and this lead to the conclusion that the CI group showed a *clear disturbance* (p.9) in motor sequencing. Moreover, when excluding outliers, the finger tapping scores correlated with language outcomes (the correlation analysis involved children with CI only). I might agree with the Authors that this last result highlights the relation between motor and language development. However, I doubt that the first result should indicate a *clear disturbance* in motor sequencing skills and that the primary cause for slower times in tapping the fingers against the thumb using the non-dominant hand should reside in auditory deprivation at birth. Instead, it might be reasonable to think that the cause (or one of the causes) leading to motor and language impairments might reside in some aspect that is shared between motor actions and language, for example predictive coding. This was the hypothesis that lead me to study 3.

4.2 Study 3: warning - imperative

The broad goal of the present study was to investigate the role of hearing experience in developing sequential timing behavior. As we saw in the previous chapters, a prediction of the ASH is that the development of timing behavior should be negatively affected by a delay of hearing exposure or, to a greater extent, by the lack of hearing exposure. This second scenario was challenged by the study of Iversen et al. (2015), who asked hearing and deaf adults to take part in a tapping task. In a tapping task, participants are usually asked to tap their index finger following an isochronous rhythmic pattern. In Iversen et al. study, three types of stimuli were used: a white square flashing on the screen, a bouncing basketball moving realistically on the screen, and an auditory tone. Deaf adults – and precisely 23 Deaf ASL signers born deaf (N=20) or became deaf before the age of three (N=3) – performed better than hearing non signers matched in age and education when synchronizing to the white flash. This result indicates that hearing experience is not needed to perform a temporal task with regular visual stimuli. On the contrary, synchronization to a static visual regular pattern resulted enhanced in Deaf individuals, presumably as a consequence of cross-modal plasticity. As a conclusion, Iversen et al. proposed that the ASH might be limited to sequencing ability, without involving timing behavior.

Still, it might be that participants of Iversen et al. (2015) did benefit of their lifelong experience with sign language, and that individuals who are born deaf in hearing families and who have no access to sign language might indeed show an impairment in transition and timing knowledge. In fact, it has been proved that sign languages exhibit rhythmic patterns (Boeys Braem, 1999) and that infants acquiring a sign language are sensitive to the rhythmic pattern of the sign language they are acquiring (Petitto et al., 2001, 2004). And even when a deaf child is not exposed to a sign language from birth, from the moment she will come into contact with a sign language she should be able to process all the characterizing features of that language, rhythmic pattern included. Considering that a characterizing feature of rhythm is the regular occurrence of stimuli across time and space, it might not be surprising to observe good temporal abilities with visual stimuli in deaf signers. On the contrary, the situation of deaf children with CI learning a spoken language is quite different.

The period of language deprivation is a direct consequence of the impossibility to perceive speech sounds. CI devices, which basically allow the delivery of electrical pulses to stimulate the primary auditory nerves via an electrode array surgically inserted in the cochlea (Hopyan-Misakyan et al., 2009), provide high levels information, even though some aspects of the speech signal are lost. For example, perception of harmonic pitch is impaired, resulting in a weak transmission of prosodic information (Chatterjee et al., 2015). A delayed exposure to speech sounds followed by good but not full perception of speech sounds might result in impairments in the encoding of transition and timing information, which might be one of the causes for subsequent language delays.

The main goal of the present study was to investigate if deaf children with CI are able to construct the abstract representation of regular isochronous stimuli. If participants can build an abstract representation of a timing sequence, they should also be able to perform actions in time with the stimuli and, given an event, anticipating the regular forthcoming stimulus. To assess those abilities, we used a warning-imperative task (Walter et al., 1964; Pagliarini, 2015). Given a regular sequence of beats, the warning-imperative task assesses participants ability to perform an action on an imperative beat after the occurrence of a warning beat. If early auditory experience is necessary for the development of this ability, deaf children with CI should be impaired compared to their hearing peers because of the delay in sound exposure.

4.2.1 Methods

Participants A group of 11 children with severe/profound hearing impairments with CI (mean age = 96.55 months, SD= 25.54, range 59 – 140) took part in this study. They were all born in hearing families and received CI at the Ospedale di Piacenza (Hospital of Piacenza, Italy) between 7 and 39 months of age (M = 21.18, SD = 11.87). Five of them had a bilateral implant and six a unilateral implant (three children used a contralateral hearing aid on the other ear and three children did not). All children with CI underwent auditory-verbal therapy, with no use of sign language or total communication. According to the neuropsychological evaluation carried out at the hospital, none of them had any mental deficit. At home, all children were exposed to one language: Italian. No child was exposed to any second language, either signed or spoken. Table 4.1 briefly summarizes biographical characteristics of children with CI.

All participants were tested individually at the Hospital of Piacenza in a single session of about 45 minutes. Informed consent was obtained from the parents of all participating children. The Hospital of Piacenza and the University of Milano-Bicocca Ethical Committees approved the study. Participants received no money for their participation.

Table 4.1: Children with CI: biographical information

ID	Etiology	Age	Age IC	IC type	PTA pre	PTA post
01	cytomegalovirus	102	27	U	120	30
02	genetic	108	17	U + HA	125	30
03	connexin 26	140	12	B	125	35
04	connexin 26	112	14	B	125	30
05	connexin 26	104	8	B	125	30
06	connexin 26	67	7	U	110	30
07	genetic	59	36	B	110	35
08	connexin 26	88	38	U	115	30
09	unknown	92	17	U + HA	90	35
10	unknown	65	18	B	105	30
11	unknown	125	39	U + HA	115	35

Age and Age IC: in months
CI type: U=unilateral, B=bilateral, HA=with contralateral HA
PTA = Pure Tone Average; pre surgey and post surgey

All children underwent an assessment battery that measured lexical and grammatical comprehension skills. The battery included the following tests:

Peabody Picture Vocabulary Test - Revised, Italian version (PPVT; Stella et al., 2000). This task examines Italian word comprehension. The examiner reads a word and the child has to choose the target picture among four alternatives.

Batteria per la valutazione del linguaggio in bambini dai 4 ai 12 anni (BVL; Marini et al., 2015 - Grammar comprehension subtest). This task examines Italian comprehension at the sentence level. The examiner reads a sentence and the child has to choose the target picture among four alternatives.

Table 4.2 briefly summarizes assessment scores.

Table 4.2: Children with CI: assessment scores

ID	PPVT - RAW	PPVT - STAND*	BVL - RAW	BVL Z-SCORE*
01	76	74	29	-2.12
02	119	105	35	-0.71
03	126	92	38	-0.08
04	104	94	37	-0.06
05	129	94	36	-0.22
06	80	91	33	0.59
07	29	70	22	-0.10
08	77	79	29	-1.78
09	89	84	36	0.17
10	71	85	25	-0.56
11	112	81	31	-2.48

*Based on normative data

Materials Participant's ability to synchronize to a regular sound pattern (that reflects the ability to predict the timing of forthcoming stimuli) was assessed using a reduced version of Pagliarini's warning-imperative task (Pagliarini, 2015, chapter 6 and 7. For further information about the warning-imperative paradigm see Walter et al., 1964). In a first habituation phase, lasting about 30 seconds, the target pattern was reproduced with a single sound, a pure tone, that I will call base tone. During the testing phase, the warning beat and the imperative beat were introduced. Warning beats and imperative beats were pairs of adjacent tones that differed from the base tone. Because the imperative beat immediately followed the warning beat, the warning beat predicted the timing of the imperative beat. On the contrary, the warning beat was not predictable. Participant's task was to click the left mouse button in synchrony with the imperative beat.

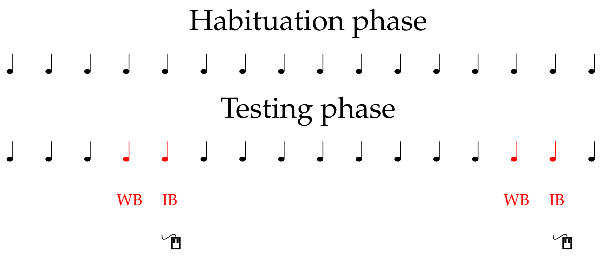


Figure 4.1: Warning Imperative task: procedure.

Participant's task was to click the left mouse button in response to the imperative beat. WB= warning beat; IB= imperative beat; 🖱= expected click

All participants were exposed to two different pattern conditions: one was regular (predictable) and one irregular (unpredictable). The two conditions corresponded to Pagliarini's rhythm 1 and 4: The regular pattern, a cadence, "had a reference tempo of 80 bpm. The beats were 440 Hz pure tones with 8 ms rise and fall times and 200 ms steady-state duration. Beats were presented with onset-to-onset intervals of 750 ms. Sequences of tones were presented in trains of 6000 ms of duration and each train contained 6 basic tones and one warning beat-imperative beat couple tones" (Pagliarini, 2015, p. 127) whereas in the irregular condition "auditory stimuli were 440 Hz pure tones with 8 ms rise and fall times and 200 ms steady-state duration. Sounds were presented with a mean onset-to-onset intervals of 750 ms ± a random error of 30% of the reference duration of 750 ms. Sequences of tones were presented in trains of 6000 ms, and each train contained 6 basic tones and one warning beat-imperative beat couple tones" (Pagliarini, 2015, p. 128). If participants could synchronize to the regular rhythm, and generate accurate predictions about the forthcoming stimuli, they should be able to click exactly in correspondence to the IB. The irregular rhythm condition served as control: the timing was unpredictable, therefore participants were not expected to generate accurate predictions.

The warning-imperative task was presented using MATLAB (Matlab, 2016, b) and Psychtoolbox Version 3.0.11 (Kleiner et al., 2007). Sounds were generated by MATLAB and played via loudspeakers.

4.2.2 Results and analysis

Errors and outliers were replaced by the median calculated on the remaining values. Errors (3%) were response given either before the warning beat or after the beat following the imperative beat. To detect outliers (12%), we used the Median Absolute Deviation (MAD) of the 10 repetitions of each rhythm (Hoaglin et al., 1983). Data point that were more than 2.5 MAD from the median were considered as outliers.

By subject results are reported in Table 4.3. As example, we depicted the performance of two children with CI in Figure 4.2.

Table 4.3: Warning - imperative task: by subject results

	Regular condition			Irregular condition		
ID	abs mean	mean	SD	abs mean	mean	SD
01	36.55	-28.95	30.54	304.00	172.00	357.97
02	429.30	429.30	93.12	307.30	303.30	127.31
03	78.20	78.20	55.21	78.90	50.30	94.83
04	87.50	87.50	52.16	236.30	187.30	210.40
05	158.80	-20.40	190.75	212.00	0.60	238.71
06	217.60	217.60	62.43	249.80	249.80	100.17
07	350.10	340.10	210.72	184.65	95.95	251.07
08	36.30	-32.90	29.42	125.90	-4.90	172.42
09	162.10	97.30	161.15	292.90	2.90	329.79
10	236.75	-236.75	124.03	225.30	-186.50	190.19
11	243.65	-243.65	103.00	328.30	-129.90	380.78

abs mean = by subject mean of the absolute value of the synchronization error;
mean = by subject mean of the the synchronization error;
SD = individual standard deviations of the synchronization error.
All values are in milliseconds.

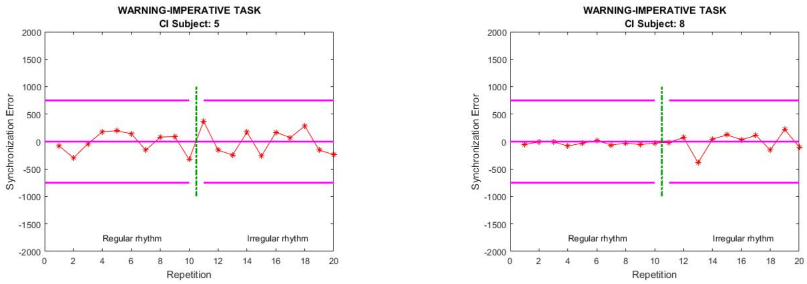


Figure 4.2: Performance of two children with CI in the warning-imperative task.

Synchronization error in ms is reported on the y axis and repetition on the x axis (1-10: regular rhythm; 11-20: irregular rhythm).

We first analyzed participants' consistency (precision) considering individual standard deviations of the synchronization error for the regular and irregular conditions. We compared the two conditions with a linear regression analysis with standard deviations as dependent variable and condition (regular vs. irregular) as predictor. Participants' standard deviations in the regular condition were smaller than standard deviations in the irregular condition (t=-3.41, p=.003, R^2=.37), indicating that participants were more precise in the regular condition than in the irregular one (see Figure 4.2 and compare repetitions 1-10 -regular condition- with repetitions 11-20 -irregular condition-).

To evaluate the ability to synchronize, we considered the absolute value of the synchronization error, i.e. the absolute value of the time difference between participant's response and the exact beginning of the imperative beat. The closest to zero was this measure, the more accurate was the predictive ability.

Given the limited sample size and the large variability that characterize the present data (and that usually characterizes performance of children with CI, see Volpato and Vernice 2014), participants' performance was evaluated individually. We considered as reference point the performance of two groups of typically developing hearing children. One group was composed by 6 years old children, whereas the second group by 10 years old children. For each child with CI, we evaluated if the absolute value of the synchronization error was age appropriate[1] by means of inverse z-points. Given the unpredictable nature of the irregular condition, we were expecting no extreme z-values (i.e. z value <-2). An extreme z-value in the irregular condition should indicate a general problem with the task. On the contrary, in case of a specific deficit in generating temporal prediction on the basis of a predictable rhythm, we should

---

[1] For each child, we considered one control group between the two reported in Appendix B, depending on the child age.

observe extreme z-values for the regular condition. Results are reported in Table 4.4.

Table 4.4: Warning - imperative task: absolute value of the synchronization error z scores

ID	Regular condition abs mean z-score	Irregular condition abs mean z-score
01	0.77	-0.46
02	-1.98	-0.48
03	0.48	0.86
04	0.42	-0.07
05	-0.08	0.07
06	-0.17	-0.02
07	-1.14	0.36
08	1.16	0.70
09	0.24	-0.27
10	-0.31	0.12
11	-0.68	-0.60

Z-scores calculated considering
normative data reported in Appendix B

As for the regular condition, z-scores ranged between a minimum of -1.98 to a maximum of 1.16, whereas for the irregular condition between -0.60 and 0.86. As results on the irregular condition showed, deaf children could comply with the task requirements and they performed very similar to hearing controls. Moreover, the fact that no child obtained a z-score lower than -2.0 in the regular condition indicates that deaf children were not impaired in the warning-imperative task.

We performed a correlation analysis between the absolute value of the synchronization error for the regular condition, the linguistic assessment (lexical and grammatical assessment) and the biographical characteristics (age, age of implantation). No correlation reached significance.

Importantly, also the correlation between the absolute value of the synchronization error and pre and post implant pure tone average was not significant. This result is important to exclude any effect due to a reduced sense of hearing.

4.2.3 Discussion

In this study, we investigated if deaf children with CI are able to construct the abstract representation of an auditory regular pattern, showing the ability to encode sequences timing. To do so, we employed a warning-imperative paradigm with a regular, predictable cadence and an irregular, unpredictable sounds pattern. Firstly, we showed that deaf children with CI were more precise in the regular compared to the irregular condition. Then, we analyzed if the absolute value of the synchronization error was age appropriate by means of inverse z-scores. Z-scores for the irregular condition fell between -1 and 1, which indicates that children with CI understood and could perform the task. Z-scores distribution for the regular condition should reflect the variability in the ability to synchronize to a regular sounds pattern and generate predictions accordingly. The vast majority of children performed within standard range (see Table 4.4), arguing against a possible deficit in encoding transition and timing of sequences. Even if we need to increase the sample size, our results suggest that deaf children with CI perform better than individuals with dyslexia in the present task. In fact, dyslexic children and adults behave different than non-dyslexic controls in the warning-imperative task (Pagliarini, 2015).

To perform the task, which required to click a mouse in synchrony with the imperative beat, fine motor abilities were fundamental. Specifically, considering the sequential nature of the task, the present results are in contrast with the disturbance on motor sequencing reported by Conway et al. (2011a). However, the present sample of children with CI performed the task with their preferred hand, and the greatest disturbances that Conway et al. found in their participants involved the non-dominant hand. Further research should analyze this issue in greater detail. Nevertheless, the present results did not meet the prediction of the ASH, i.e. this small group of children with CI showed the ability to encode sequences timing, suggesting that a short period of auditory and language deprivation *per se* does not affect the developing of this skill. In the introduction of this chapter, I presented the argument that deaf children with CI might be impaired in encoding transition and timing information because their exposure to speech and language is not only delayed, but also slightly partial, from the phonetic point of view. Even if this argument looks sound to me, the fact that our brain emits a mismatch negativity also with non-auditory stimuli (e.g. Pause and Krauel, 2000; Pazo-Alvarez et al., 2003) makes the result of the present study not surprising. As Conway et al. (2011a) stressed, *"the brain is an integrated functional system"* (p.2). According to Conway et al., the main consequence of this is that *"it is likely that a period of auditory deprivation occurring early in development may have secondary cognitive and neural sequelae in addition to the obvious hearing-related effects"* (p.2). From my point of view, it is in the same way likely that the intact sensory systems might play a crucial role in developing those cognitive skills that are sustained (probably to a greater extent) by a damaged sensory system.

I will now focus on the relationship between transition and timing knowledge and language skill. We did not find any correlation between the synchronization ability and the linguistic scores. A first explanation might reside in the small sample size. It is reasonable to think that by increasing the number of participants it might be possible to observe a significant correlation between the synchronization error and the linguistic measures. Another possibility would be to add further linguistic measures, assessing not only comprehension, but production too. A third consideration is that Conway et al. (2011a) found a positive correlation between fine motor abilities and linguistic measures in a group of deaf children with CI with lower linguistic skill compared to hearing peers. On the contrary, the majority of our participants was not impaired considering language comprehension skills (see Table 4.2).

In addition to the small sample size, a limit of the present study is the great variability that characterizes the performance not only of children with CI, but also of the normative sample (see Appendix B). As a matter of fact, the task was difficult, especially for the younger children. In the future, I might assess children with CI aged less than 7 years with a standard and more simple tapping task as the one used by Iversen et al. (2015).

It might be interesting to continue using the present paradigm restricting the participation to older children. Moreover, instead of using only a regular cadence, it might be preferable to assess also the ability to synchronize to a regular rhythmic pattern, as Pagliarini (2015) did in testing dyslexic children and adults. In fact, as Lashley (1951) argued (see also Fitch and Martins, 2014), one of the crucial building blocks of language, music (rhythm) and complex actions is hierarchy, and hierarchy is observable in complex rhythmic patterns, but not in isochronous rhythms (like the cadence that we used in the present task).

To sum up, the aim of the present study was to test the ASH by analyzing transition and timing skill in deaf children with CI. All in all, results suggested that deaf children with CI can develop good temporal skill, raising further doubts on the validity of the ASH.

CHAPTER 5

General conclusions and future directions

In this final chapter, I go back to the auditory scaffolding hypothesis (ASH) and I further explain why this hypothesis needs to be revisited. My final discussion on the ASH substantially benefits from the experimental results reported in chapters 2 - 4 and from the recent published works arguing against the ASH (Iversen et al., 2015; Hall et al., 2017; von Koss Torkildsen et al., 2018). Then, I sum up the strengths and weakness of my dissertation.

5.1 Converging evidence against the ASH

5.1.1 The case of deaf children with cochlear implant

As reviewed multiple times during this dissertation, the ASH states that sound perception builds "an auditory scaffolding for time and serial-order behavior". As a corollary to this, "under conditions of auditory deprivation, auditory scaffolding is absent, resulting in neural reorganization and a disturbance to cognitive sequencing abilities" (Conway et al., 2009, p.275). One of the main evidence in favor of this hypothesis came from a study showing that deaf children with CI were impaired in an implicit sequence learning task (Conway et al., 2011b). As I explained in chapter 1.2, the first concern about this hypothesis might be that data from a population of deaf children with cochlear implant (CI) shouldn't be considered appropriate to make inference about the entire deaf population.

A different hypothesis, called by Hall et al. (2017) Language Scaffolding Hypothesis, suggests that the implicit sequence learning impairments in children with CI reported in Conway et al. (2011b) might be a consequence of a period of language deprivation, instead of a period of auditory deprivation. This was also my initial hypothesis. This idea was based on the assumption that deaf children with CI do show disturbances in time and serial-order behavior. It turned out that this assumption was not entirely correct. Deaf children with CI, at least if implanted early (i.e. by age 4), seem to be able to perform tasks assessing serial-order skills (Hall et al., 2017; von Koss Torkildsen et al., 2018) and synchronization ability (chapter 4) with the same variability that characterize typical developing children. As a further consideration, the ASH would predict broad language impairments in deaf children, especially in the syntax domain, where sequencing is central. On the contrary, fine linguistic analysis on language impairments in deaf children indicated that some syntactic structures with non-canonical order of constituents (where sequencing is crucial) are spared (e.g. passive sentences) whereas others are impaired (e.g. wh-questions) (Ruigendijk and Friedmann, 2017). Therefore, there is no evidence for a generalized "sequencing disability".

With regard to the possibility of developing good sequencing skill in the context of delayed exposure to sound and speech, one should also consider that research suggests that critical periods in deaf children may stay open more, and at least until 3.5 years (see Werker and Hensch, 2015, for a review).

5.1.2 The case of Deaf signers

A fundamental piece of the puzzle for the assessment of the ASH needs to come from severe/profound Deaf signers behavior. In fact, if the auditory scaffolding was so central, the greatest problems should occur in those people who live in the absence of hearing experience. Currently (at least where I have collected data, i.e. in Italy and in California), this is the case of the vast majority of Deaf signers belonging to a Deaf community. The experimental results reported in chapter 2 and chapter 3 showed that Deaf signers' results on AGL and statistical learning tasks somehow differ from hearing non signers' results. In fact, the relation between AGL/statistical learning measures and cognitive (memory) and linguistic (reading) measures was characterized by peculiar patterns comparing Deaf and hearing participants. Those results are intriguing and need further investigation. Nevertheless, Deaf signers did not show sequence learning impairments, neither at the level of chunking contiguous items, not at the level of extracting abstract rules from sequential input. In line with those adult studies, Hall et al. (2017) showed good statistical learning abilities in Deaf children. As a matter of fact, it is well known that language acquisition and related critical periods do not depend on modality, i.e. language development of a visual language follows the same pattern than that of a spoken language (see e.g. Mayberry et al., 2002). Being sequential skills so important for language, the possibility to acquire a fully-fledged language in absence of auditory stimuli by itself introduces a potential challenge for the ASH.

The ASH at different levels of sequences representation

As described in the first chapter of this thesis, the main goal of my work was to investigate the ASH in light of Dehaene Taxonomy (Dehaene et al., 2015). Table 5.1 provides a summary of experimental findings about sequencing abilities in Deaf individuals, divided according to Dehaene Taxonomy. In this table, I focus only on Deaf signers without any kind of hearing experience because I think that it is the best population in order to test the ASH.

As Table 5.1 indicates, Deaf participants do not show impairments in sequencing behavior. The only case in which Deaf individuals performed worse than hearing individuals is the case of the finite state grammar of chapter 2. A speculation could be that, in the AGL task, Deaf participants could not generalize the rule of the finite state grammar because they focused on the statistical properties of the input. In fact, there was an important difference between the familiarization strings of the finite-state grammar with respect to the supra regular grammars. In the finite state grammar, participants were exposed to three types of sequences, ABBA, ABBBA and ABBBBBA. In the supra-regular grammars, participants were exposed to many types of sequences (i.e., considering the mirror grammar: AB.BA as well as BA.AB, ABB.BBA as well as BAA.AAB, AAB.BAA, BBA.ABB, etc.). Taking into account the statistical properties of the input when the input consists of three type of strings (not the exact three types of sequences because A and B corresponded to categories, but elements inside a category were very similar, see Figure 2.8) might be a reasonable strategy leading to rejecting grammatical extensions, thus failing the task.

5.2 An extended framework for understanding the relations among sound, cognitive sequencing, and language development

In chapter 1 I described the framework proposed by Conway et al. (2009) to explain how sound, cognitive sequencing and spoken language relates

Table 5.1: Performance of Deaf people in tasks tapping different levels of sequences representation.

Level	Participants	Evidence	Reference
Transition and timing	Deaf adults	better than hearing individuals when synchronizing with visual flashes	Iversen et al. (2015)
Chunking	Deaf adults	visual statistical learning skills: same as hearing controls	Chapter 3
Ordinal knowledge	Deaf adults	visual statistical learning skills: same as hearing controls	Chapter 3
	Deaf children	evidence of sequence learning using a serial reaction time task	Hall et al. (2017)
Algebraic patterns	Deaf adults	only a minority learned a finite state grammar	Chapter 2
Nested tree structures	Deaf adults	the majority learned mirror and copy grammars	Chapter 2

with each other (see Figure 1.2). In light of the findings presented in this dissertation and considering the recent publications about the ASH, that framework needs to be revised. I do not argue that that framework was wrong, I simply argue that it was incomplete. In fact, timing and sequencing skills seem to develop even in the absence of sound experience. Therefore, further input should sustain their development. Visual input (perception of movement, perception of repetitions across space...) might be a piece of this input (Figure 5.1). An interesting line of research to determine what kind of input is sufficient to develop timing and sequencing skills might be the assessment of those abilities in individuals with congenital/early acquired deafblindness. However, considering that the incidence of deaf-blindness is very low and that it is common for children with deaf-blindness to have associated disabilities (more than 90%, as reported in Killoran 2007), pursue this line of research might be unrealistic.

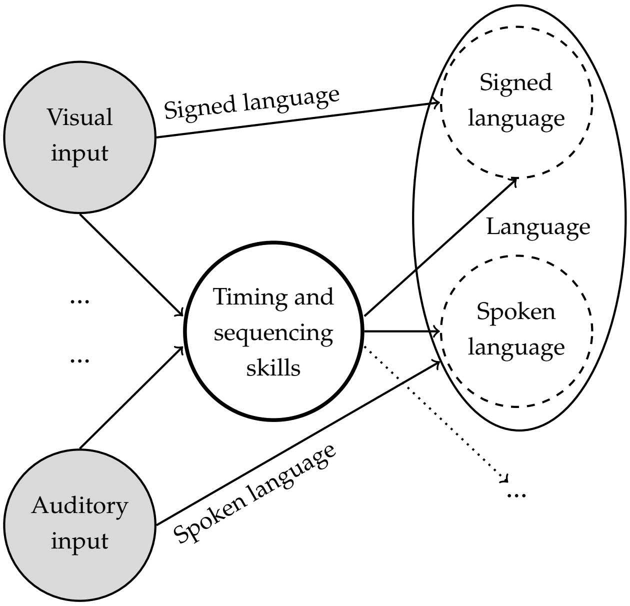


Figure 5.1: An extended scaffolding for timing and sequencing skills
Timing and sequencing skills correspond to the five levels of sequences encoding described in chapter 1. Humans can develop timing and sequencing skills thanks to inputs of at least two different modalities, visual and auditory. Timing and sequencing skills are among the building blocks of language, but also of further high-level cognitive functions.

Figure 5.1 represents my proposal for an extended framework for understanding the relations among sound, cognitive sequencing, and language development. For the sake of simplicity, all levels of timing and sequencing skills were represented in a single node. This node contains a set of cognitive abilities that are among the building blocks of higher levels cognitive functions, e.g. language. The development of timing and sequencing skills is sustained by external input, both auditory and visual. (The role of further modalities needs additional investigations.) Timing and sequencing skills are constrained by modality as well as language is constrained by modality, however what I argue is that hearing is not strictly required for those skills to develop (as well as hearing is not strictly required for language to develop).

5.3 Strengths and weakness of this dissertation

In this dissertation I collected three different works aiming at investigating timing and sequencing abilities in deaf individuals. Having considered timing and sequencing skills at different level of complexity was the major strength of the present work. However, I considered different groups of people, which constitutes a weakness. In fact, also considering the studies reported in chapter 2 and 3 with Deaf adults as participants, having three different groups makes the comparison of performance across tasks very difficult.

A further weakness, especially considering chapter 2 and 3 is that all Deaf participants were bilingual, mastering both a sign language (either LIS or ASL) and the written form of a spoken language (Ttalian or American English), whereas the great majority of hearing participants were monolingual.

Considering chapter 4, as I have extensively discussed in section 4.2.3, the major weakness consists of sample size, which need to be incremented, especially considering the high variability that characterize performances of deaf children with CI.

Nevertheless, the data I collected together converge to the evidence that no auditory scaffolding is needed for timing and sequencing skills to develop.

APPENDIX A

Further information on the VAGL task

A.1 Experiment 1-4: example of grammatical stimuli

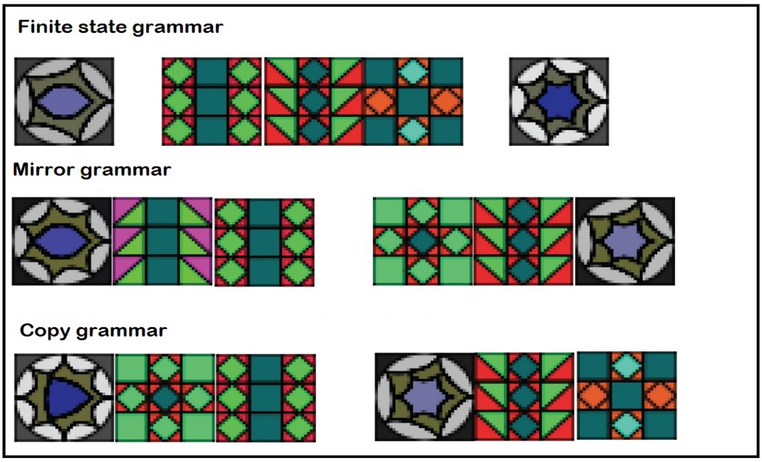


Figure A.1: Examples of stimuli sequences for all three grammars with N=3

A.2 Experiment 4: summary of test stimuli

Table A.1: Description of test stimuli

		Grammar	
Condition	Finite State	Mirror	Copy
Correct N=2,3	20	20	20
Correct N=4	10	10	10
Correct N=6	6	6	6
Missing tile N=4	10	10	10
Missing tile N=6	6	6	6
Wrong tile N=4	10	10	20
Missing tile N=2,3	10	10	10
Wrong tile N=2,3	15	15	15
Total	87	87	87

Categories used in Tables 2.11 and 2.12

Correct N=2,3 : Grammatical

Correct N=4 & Correct N=6 : Grammatical extension

Missing tile N=2,3 & Missing tile N=4 & Missing tile N=6 : Foil missing

Wrong tile N=2,3 & Wrong tile N=6 : Wrong tile

APPENDIX B

Warning - Imperative task

B.1 Control data. Group Hearing-6

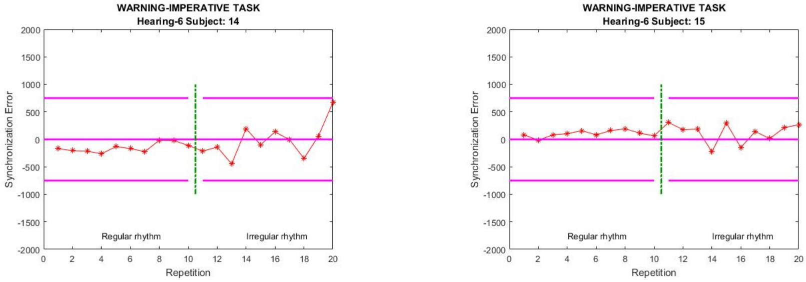


Figure B.1: Performance of two 6 years old hearing children in the warning-imperative task.

Synchronization error in ms is reported on the y axis and repetition on the x axis (1-10: regular rhythm; 11-20: irregular rhythm).

Group Hearing-6 was composed by 17 hearing children aged between 5 and 7 years (mean age: 6).

Figure B.1 shows the performance of two hearing children aged 6.

Repetitions 1-10 correspond to the regular rhythm and repetitions 11-20 to the irregular rhythm. Group performance (mean and standard deviations) is reported in Table B.1.

Table B.1: Group Hearing-6 - performance

	Mean	SD
Regular rhythm		
Synchronization error	80.90	224.27
Absolute value of synchronization error	194.85	136.71
Irregular rhythm		
Synchronization error	121.08	275.51
Absolute value of synchronization error	246.22	172.25

B.2 Control data. Group Hearing-10

Group Hearing-10 was composed by 29 hearing children aged between 9 and 11 years (mean age: 10).

Typical performance of two Hearing-10 children is depicted in Figure B.2. Group performance (mean and standard deviations) is reported in Table B.2.

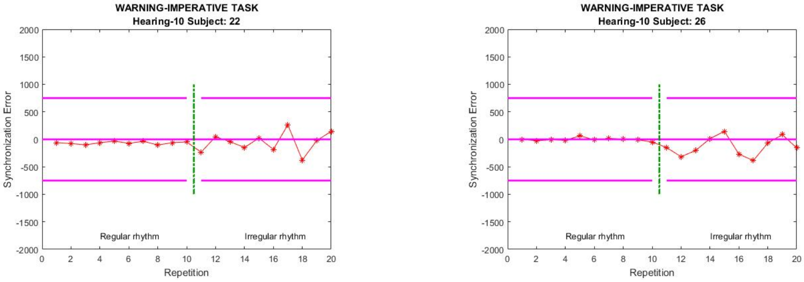


Figure B.2: Performance of two 10 years old hearing children in the warning-imperative task.

Synchronization error in ms is reported on the y axis and repetition on the x axis (1-10: regular rhythm; 11-20: irregular rhythm).

Table B.2: Group Hearing-10 - performance

	Mean	SD
Regular rhythm		
Synchronization error	-48.40	199.10
Absolute value of synchronization error	146.91	142.60
Irregular rhythm		
Synchronization error	-33.63	281.61
Absolute value of synchronization error	224.83	172.37

Bibliography

Acheson, D. J., Wells, J. B., and MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods*, 40(1):278–289.

Andrews, S. and Hersch, J. (2010). Lexical precision in skilled readers: Individual differences in masked neighbor priming. *Journal of Experimental Psychology: General*, 139(2):299.

Arciuli, J. and Simpson, I. C. (2011). Statistical learning in typically developing children: the role of age and speed of stimulus presentation. *Developmental Science*, 14(3):464–473.

Arciuli, J. and Simpson, I. C. (2012). Statistical learning is related to reading ability in children and adults. *Cognitive Science*, 36(2):286–304.

Bach, E., Brown, C., and Marslen-Wilson, W. (1986). Crossed and nested dependencies in german and dutch: A psycholinguistic study. *Language and Cognitive Processes*, 1(4):249–262.

Basso, A., Capitani, E., and Laiacona, M. (1987). Raven's coloured progressive matrices: normative values on 305 adult normal controls. *Functional Neurology*, 2(2):189–194.

Bates, E. and Elman, J. (1996). Learning rediscovered. *Science*, 274(5294):1849.

Bavelier, D., Newport, E. L., Hall, M., Supalla, T., and Boutla, M. (2008). Ordered short-term memory differs in signers and speakers: Implications for models of short-term memory. *Cognition*, 107(2):433–459.

Boeys Braem, P. (1999). Rhythmic temporal patterns in the signing of deaf early and late learners of swiss german sign language. *Language and Speech*, 42(2-3):177–208.

Bolognini, N., Cecchetto, C., Geraci, C., Maravita, A., Pascual-Leone, A., and Papagno, C. (2012). Hearing shapes our perception of time: temporal discrimination of tactile stimuli in deaf people. *Journal of Cognitive Neuroscience*, 24(2):276–286.

Boutla, M., Supalla, T., Newport, E. L., and Bavelier, D. (2004). Short-term memory span: insights from sign language. *Nature Neuroscience*, 7(9):997–1002.

Brentari, D. (1998). *A prosodic model of sign language phonology*. Mit Press.

Brentari, D. (2006). Effects of language modality on word segmentation: An experimental study of phonological factors in a sign language. *Papers in Laboratory Phonology*, 8:155–164.

Bruce, V. G. and Morgan, M. J. (1975). Violations of symmetry and repetition in visual patterns. *Perception*, 4(3):239–249.

Bulf, H., Johnson, S. P., and Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition*, 121(1):127–132.

Chatterjee, M., Zion, D. J., Deroche, M. L., Burianek, B. A., Limb, C. J., Goren, A. P., Kulkarni, A. M., and Christensen, J. A. (2015). Voice emotion recognition by cochlear-implanted children and their normally-hearing peers. *Hearing Research*, 322:151–162.

Chesi, C. and Moro, A. (2014). Computational complexity in the brain. In Newmeyer, F. J. and Preston, L. B., editors, *Measuring grammatical complexity*, pages 264–280. Oxford University Press.

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory*, 2(3):113–124.

Chomsky, N. (1957). Syntactic structures. *Language*, 33(3 Part 1):375–408.

Chomsky, N. (1959). A review of b.f. skinner's verbal behavior. *Language*, 35(1):26–58.

Chomsky, N. (1980). Rules and representations. *Behavioral and Brain Sciences*, 3(1):1–15.

Cohen, M. (1997). Children's memory scale. san antonio: Psychological corporation.

Conway, C. M. and Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(1):24–39.

Conway, C. M., Karpicke, J., Anaya, E. M., Henning, S. C., Kronenberger, W. G., and Pisoni, D. B. (2011a). Nonverbal cognition in deaf children following cochlear implantation: Motor sequencing disturbances mediate language delays. *Developmental Neuropsychology*, 36(2):237–254.

Conway, C. M., Pisoni, D. B., Anaya, E. M., Karpicke, J., and Henning, S. C. (2011b). Implicit sequence learning in deaf children with cochlear implants. *Developmental Science*, 14(1):69–82.

Conway, C. M., Pisoni, D. B., and Kronenberger, W. G. (2009). The importance of sound for cognitive sequencing abilities the auditory scaffolding hypothesis. *Current Directions in Psychological Science*, 18(5):275–279.

Corsi, P. M. and Michael, P. (1972). *Human memory and the medial temporal region of the brain*, volume 34. McGill University Montreal.

Crasborn, O. A., Van Der Kooij, E., Waters, D., Woll, B., and Mesch, J. (2008). Frequency distribution and spreading behavior of different types of mouth actions in three sign languages. *Sign Language & Linguistics*, 11(1):45–67.

Culy, C. (1985). The complexity of the vocabulary of bambara. *Linguistics and Philosophy*, 8(3):345–351.

De Vries, M. H., Barth, A. C., Maiworm, S., Knecht, S., Zwitserlood, P., and Flöel, A. (2010). Electrical stimulation of broca's area enhances implicit learning of an artificial grammar. *Journal of Cognitive Neuroscience*, 22(11):2427–2436.

Dehaene, S., Meyniel, F., Wacongne, C., Wang, L., and Pallier, C. (2015). The neural representation of sequences: from transition probabilities to algebraic patterns and linguistic trees. *Neuron*, 88(1):2–19.

Evans, J. L., Saffran, J. R., and Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 52(2):321–335.

Fiser, J. and Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, 12(6):499–504.

Fitch, W. and Martins, M. D. (2014). Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences*, 1316(1):87–104.

Fitch, W. T., Friederici, A. D., and Hagoort, P. (2012). Pattern perception and computational complexity: introduction to the special issue. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1598):1925.

Fitch, W. T. and Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, 303(5656):377–380.

Frost, R., Armstrong, B. C., Siegelman, N., and Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, 19(3):117–125.

Frost, R., Siegelman, N., Narkiss, A., and Afek, L. (2013). What predicts successful literacy acquisition in a second language? *Psychological Science*, 24(7):1243–1252.

Gabay, Y., Thiessen, E. D., and Holt, L. L. (2015). Impaired statistical learning in developmental dyslexia. *Journal of Speech, Language, and Hearing Research*, 58(3):934–945.

Gebhart, A. L., Newport, E. L., and Aslin, R. N. (2009). Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. *Psychonomic Bulletin & Review*, 16(3):486–490.

Gentner, T. Q., Fenn, K. M., Margoliash, D., and Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature*, 440(7088):1204.

Giustolisi, B., Mereghetti, E., and Cecchetto, C. (2017). Phonological blending or code mixing? why mouthing is not a core component of sign language grammar. *Natural Language & Linguistic Theory*, 35(2):347–365.

Giustolisi, B., Vergallito, A., Cecchetto, C., Varoli, E., and Romero Lauro, L. J. (2018). Anodal transcranial direct current stimulation over left inferior frontal gyrus enhances sentence comprehension. *Brain and Language*, 176:36–41.

Goldin-Meadow, S. and Mayberry, R. I. (2001). How do profoundly deaf children learn to read? *Learning Disabilities Research & Practice*, 16(4):222–229.

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

Gomez, R. L. and Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, 70(2):109–135.

Hall, M. L., Eigsti, I.-M., Bortfeld, H., and Lillo-Martin, D. (2017). Auditory access, language access, and implicit sequence learning in deaf children. *Developmental Science*.

Hauser, P. C., Lukomski, J., and Hillman, T. (2008). Development of deaf and hard-of-hearing students' executive function. In Marschark, M. and Hauser, P. C., editors, *Deaf cognition: Foundations and outcomes*, chapter 1, pages 286–308. Oxford University Press, Oxford.

Hauser, P. C., Paludneviciene, R., Riddle, W., Kurz, K. B., Emmorey, K., and Contreras, J. (2015). American sign language comprehension test: A tool for sign language researchers. *Journal of Deaf Studies and Deaf Education*, 21(1):64–69.

Hoaglin, D. C., Mosteller, F., and Tukey, J. W. (1983). Understanding robust and exploratory data anlysis. *Wiley Series in Probability and Mathematical Statistics, New York: Wiley, 1983, edited by Hoaglin, David C.; Mosteller, Frederick; Tukey, John W.*

Hopyan-Misakyan, T. M., Gordon, K. A., Dennis, M., and Papsin, B. C. (2009). Recognition of affective speech prosody and facial affect in deaf children with unilateral right cochlear implants. *Child Neuropsychology*, 15(2):136–146.

Howard, J. H., Howard, D. V., Japikse, K. C., and Eden, G. F. (2006). Dyslexics are impaired on implicit higher-order sequence learning, but not on implicit spatial context learning. *Neuropsychologia*, 44(7):1131–1144.

Huybregts, R. (1984). The weak inadequacy of context-free phrase structure grammars. In *Van periferie naar kern*, pages 81–99. Foris.

Iversen, J. R., Patel, A. D., Nicodemus, B., and Emmorey, K. (2015). Synchronization to auditory and visual rhythms in hearing and deaf individuals. *Cognition*, 134:232–244.

Jäger, G. and Rogers, J. (2012). Formal language theory: refining the chomsky hierarchy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1598):1956–1970.

Joshi, A., Vijay-Shanker, K., and Weir, D. (1991). The convergence of mildly context-sensitive grammar formalisms. Technical Report 539, University of Pennsylvania.

Joshi, A. K. (1990). Processing crossed and nested dependencies: An automation perspective on the psycholinguistic results. *Language and Cognitive Processes*, 5(1):1–27.

Kaufman, A. S. and Kaufman, N. L. (2004). *Kaufman brief intelligence test*. Wiley Online Library.

Kelly, S. W., Griffiths, S., and Frith, U. (2002). Evidence for implicit sequence learning in dyslexia. *Dyslexia*, 8(1):43–52.

Killoran, J. (2007). The national deafblind child count: 1998–2005 in review. *National Technical Assistance Consortium For Children and Young Adults who are Deaf-Blind*.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C., et al. (2007). What's new in psychtoolbox-3. *Perception*, 36(14):1.

Korkman, M., Kirk, U., and Kemp, S. (1998). *NEPSY: A Developmental Neuropsychological Assessment*. Psychological Corporation.

Kowalska, J. and Szelag, E. (2006). The effect of congenital deafness on duration judgment. *Journal of Child Psychology and Psychiatry*, 47(9):946–953.

Lashley, K. S. (1951). The problem of serial order in behavior. In Jeffress, L., editor, *Cerebral mechanisms in behavior; the Hixon Symposium*, pages 112–136. Wiley.

Lum, J. A., Conti-Ramsden, G., Morgan, A. T., and Ullman, M. T. (2014). Procedural learning deficits in specific language impairment (sli): A meta-analysis of serial reaction time task performance. *Cortex*, 51:1–10.

Marcus, G. F., Vijayan, S., Rao, S. B., and Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, 283(5398):77–80.

Marini, A., Marotta, L., Bulgheroni, S., and Fabbro, F. (2015). Batteria per la valutazione del linguaggio in bambini dai 4 ai 12 anni (bvl_4-12). *Firenze: Giunti OS*.

Markwardt, F. C. (1989). *Peabody individual achievement test-revised: PIAT-R*. American Guidance Service Circle Pines.

Marschark, M. and Hauser, P. C. (2008). Cognitive underpinnings of learning by deaf and hard-of-hearing students: Differences, diversity and directions. In Marschark, M. and Hauser, P. C., editors, *Deaf cognition: Foundations and outcomes*, chapter 1, pages 3–23. Oxford University Press, Oxford.

Mathews, R. C. (1990). Abstractness of implicit grammar knowledge: Comments on perruchet and pacteau's analysis of synthetic grammar learning. *Journal of Experimental Psycology: General*, 119(4):412–416.

Matlab, U. G. (2016). The mathworks. *Inc., Natick, MA*, b.

Mayberry, R. I., Lock, E., and Kazmi, H. (2002). Development: Linguistic ability and early language exposure. *Nature*, 417(6884):38–38.

Miller, G. A. (1958). Free recall of redundant strings of letters. *Journal of Experimental Psychology*, 56(6):485.

Miller, G. A. (1967). Project grammarama. In *Psychology of communication*. Basic Books.

Näätänen, R. (2003). Mismatch negativity: clinical research and possible applications. *International Journal of Psychophysiology*, 48(2):179–188.

Näätänen, R., Gaillard, A. W., and Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42(4):313–329.

Nissen, M. J. and Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19(1):1–32.

Novogrodsky, R., Henner, J., Caldwell-Harris, C., and Hoffmeister, R. (2017). The development of sensitivity to grammatical violations in american sign language: Native versus nonnative signers. *Language Learning*.

Pagliarini, E. (2015). *Predictive timing in developmental dyslexia: a new hypothesis. Anticipatory skills across language and motor domains*. PhD thesis, Dottorato in Psicologia sperimentale, linguistica e neuroscienze cognitive.

Pause, B. M. and Krauel, K. (2000). Chemosensory event-related potentials (cserp) as a key to the psychology of odors. *International Journal of Psychophysiology*, 36(2):105–122.

Pazo-Alvarez, P., Cadaveira, F., and Amenedo, E. (2003). Mmn in the visual modality: a review. *Biological Psychology*, 63(3):199–236.

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition*, 17(4):398–422.

Perruchet, P. and Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General*, 119(3):264.

Perruchet, P. and Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, 10(5):233–238.

Petitto, L. A., Holowka, S., Sergio, L. E., Levy, B., and Ostry, D. J. (2004). Baby hands that move to the rhythm of language: hearing babies acquiring sign languages babble silently on the hands. *Cognition*, 93(1):43–73.

Petitto, L. A., Holowka, S., Sergio, L. E., and Ostry, D. (2001). Language rhythms in baby hand movements. *Nature*, 413(6851):35.

Piattelli-Palmarini, M. (1994). Ever since language and learning: Afterthoughts on the piaget-chomsky debate. *Cognition*, 50(1):315–346.

Purves, D., Brannon, E. M., Cabeza, R., Huettel, S. A., LaBar, K. S., Platt, M. L., and Woldorff, M. G. (2008). *Principle of Cognitive Neuroscience*. Sunderland: Sinauer Associates, Inc.

Qi, S. and Mitchell, R. E. (2011). Large-scale academic achievement testing of deaf and hard-of-hearing students: Past, present, and future. *Journal of Deaf Studies and Deaf Education*, 17(1):1–18.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raven, J. (1965). Guide to using the coloured matrices, sets a, a b, b. *London: HK Lewis*.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6(6):855–863.

Repp, B. H. and Penel, A. (2002). Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology-Human Perception and Performance*, 28(5):1085–1099.

Repp, B. H. and Penel, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychological Research*, 68(4):252–270.

Ruigendijk, E. and Friedmann, N. (2017). A deficit in movement-derived sentences in german-speaking hearing-impaired children. *Frontiers in Psychology*, 8:689.

Rüsseler, J., Gerth, I., and Münte, T. F. (2006). Implicit learning is intact in adult developmental dyslexic readers: Evidence from the serial reaction time task and artificial grammar learning. *Journal of Clinical and Experimental Neuropsychology*, 28(5):808–827.

Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996a). Statistical learning by 8-month-old infants. *Science*, pages 1926–1928.

Saffran, J. R., Newport, E. L., and Aslin, R. N. (1996b). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35(4):606–621.

Shieber, S. M. (1985). Evidence against the context-freeness of natural language. In *The Formal complexity of natural language*, pages 320–334. Springer.

Siegelman, N., Bogaerts, L., and Frost, R. (2017). Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behavior Research Methods*, 49(2):418–432.

Sigurdardottir, H. M., Danielsdottir, H. B., Gudmundsdottir, M., Hjartarson, K. H., Thorarinsdottir, E. A., and Kristjánsson, Á. (2017). Problems with visual statistical learning in developmental dyslexia. *Scientific Reports*, 7(1):606.

Skinner, B. F. (1957). *Verbal behavior*. Englewood Cliffs.

Spencer, M., Kaschak, M. P., Jones, J. L., and Lonigan, C. J. (2015). Statistical learning is related to early literacy-related skills. *Reading and Writing*, 28(4):467–490.

Stella, G., Pizzoli, C., and Tressoldi, P. (2000). Ppvt-r, peabody picture vocabulary test–revised. *Omega Edizioni*.

Stobbe, N., Westphal-Fitch, G., Aust, U., and Fitch, W. T. (2012). Visual artificial grammar learning: comparative research on humans, kea (nestor notabilis) and pigeons (columba livia). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1598):1995–2006.

Supalla, T., Hauser, P. C., and Bavelier, D. (2014). Reproducing american sign language sentences: cognitive scaffolding in working memory. *Frontiers in Psychology*, 5.

ten Cate, C. and Okanoya, K. (2012). Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 367(1598):1984–1994.

Treder, M. S. (2010). Behind the looking-glass: A review on human symmetry perception. *Symmetry*, 2(3):1510–1543.

Treiman, R. and Kessler, B. (2006). Spelling as statistical learning: Using consonantal context to spell vowels. *Journal of Educational Psychology*, 98(3):642.

Tversky, A. and Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90(4):293–315.

Uddén, J., Folia, V., Forkstam, C., Ingvar, M., Fernández, G., Overeem, S., Van Elswijk, G., Hagoort, P., and Petersson, K. M. (2008). The inferior frontal cortex in artificial syntax processing: An rtms study. *Brain Research*, 1224:69–78.

Uddén, J., Ingvar, M., Hagoort, P., and Petersson, K. M. (2012). Implicit acquisition of grammars with crossed and nested non-adjacent dependencies: Investigating the push-down stack model. *Cognitive Science*, 36(6):1078–1101.

Van Heijningen, C. A., De Visser, J., Zuidema, W., and Ten Cate, C. (2009). Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proceedings of the National Academy of Sciences*, 106(48):20538–20543.

Volpato, F. and Vernice, M. (2014). The production of relative clauses by italian cochlear-implanted and hearing children. *Lingua*, 139:39–67.

von Koss Torkildsen, J., Arciuli, J., Haukedal, C. L., and Wie, O. B. (2018). Does a lack of auditory experience affect sequential learning? *Cognition*, 170:123–129.

Wacongne, C., Changeux, J.-P., and Dehaene, S. (2012). A neuronal model of predictive coding accounting for the mismatch negativity. *Journal of Neuroscience*, 32(11):3665–3678.

Wagemans, J. (1997). Characteristics and models of human symmetry detection. *Trends in Cognitive Sciences*, 1(9):346–352.

Walter, W. G., Cooper, R., Aldridge, V., McCallum, W., and Winter, A. (1964). Contingent negative variation: an electric sign of sensori-motor association and expectancy in the human brain. *Nature*, 203(4943):380–384.

Werker, J. F. and Hensch, T. K. (2015). Critical periods in speech perception: new directions. *Annual Review of Psychology*, 66.

Wilkinson, G. S. and Robertson, G. (2006). Wide range achievement test 4 (wrat4). *Psychological Assessment Resources, Inc, Lutz, FL.*

Wilson, M. and Emmorey, K. (1997). Working memory for sign language: A window into the architecture of the working memory system. *Journal of Deaf Studies and Deaf Education*, 2(3):121–130.

Woodward, J. C. (1972). Implications for sociolinguistic research among the deaf. *Sign Language Studies*, pages 1–7.

Zarco, W., Merchant, H., Prado, L., and Mendez, J. C. (2009). Subsecond timing in primates: comparison of interval production between human subjects and rhesus monkeys. *Journal of Neurophysiology*, 102(6):3191–3202.

Zuidema, L. A. (2012). The grammar workshop: Systematic language study in reading and writing contexts. *English Journal*, 101(5):63–71.