Domain-specificity in the Acquisition of Non-adjacent Dependencies

Jennifer A. Sturm

PhD

2011

Domain-specificity in the Acquisition of Non-adjacent Dependencies

Jennifer A. Sturm

A thesis submitted in partial fulfilment of the requirements of Northumbria University for the degree of Doctor of Philosophy.

This research project was carried out in the Department of Psychology, School of Life Sciences, Northumbria University.

August 2011

Abstract

Domain-specificity in the acquisition of non-adjacent dependencies

At the forefront of investigations into the cognitive underpinnings of language acquisition is the question of domain-specificity, i.e. whether the processes involved in learning language are unique to language. Recent investigations suggest that the mechanisms employed in language learning are also involved in sequential learning of non-linguistic stimuli and are therefore domain-general.

Non-adjacent dependencies are an important feature of natural languages. They describe relationships between two elements separated by an arbitrary number of intervening items, and thus potentially pose a challenge for learners. As a hallmark of natural languages they are ubiquitous, an example from English being subject-verb agreement: *The socks on the floor are red*. Here, learners are required to track the dependencies amongst the two underlined elements across an intervening prepositional phrase. Importantly, it has been shown that non-adjacent dependencies can be learned in the linguistic (Gómez, 2002) and non-linguistic (Creel, Newport & Aslin, 2004) domain.

The majority of work presented in this thesis is based on Gómez's (2002) artificial language learning experiment involving non-adjacent dependencies, adapted to directly compare adults' learning in the linguistic and non-linguistic

domain, in order to build a comprehensive map showing factors and conditions that enhance/ inhibit the learnability of non-adjacencies. Experiment 1 shows that the Gestalt Principle of Similarity is not a requirement for the detection of non-adjacent dependencies in the linguistic domain. Experiment 2 aims to explore the robustness of the ability to track non-adjacent regularities between linguistic elements by removing cues that indicate the correct level of analysis (i.e. interword breaks). Experiments 3 and 4 study domain-specificity in the acquisition of non-adjacencies, and show that non-adjacent dependencies are learnable in the linguistic and non-linguistic domain, provided that the non-linguistic materials are simple and lacking internal structure. However, language is rich in internal structure: it is combinatorial on the phonemic/ orthographic level in that it recombines elements (phonemes/ graphemes) to form larger units. When exposed to non-linguistic stimuli which capture this componential character of language, adult participants fail to detect the non-adjacencies. However, when exposed to non-componential non-linguistic materials, adult participants succeed in learning the non-adjacent dependencies. Experiment 5 looks at modality effects in the acquisition of non-adjacent dependencies across the linguistic and non-linguistic domain. Experiment 6 provides evidence that high familiarity with componential non-linguistic patterns does not result in the correct extraction of non-adjacencies in sequence learning tasks involving these patterns.

Overall, the work presented here demonstrates that the acquisition of non-adjacent dependencies is a domain-general ability, which is guided by stimulus simplicity.

Table of Contents

List of Tables

List of Figures

List of Appendices

Acknowledgments

First and foremost, I would like to thank my supervisors Dr Kenny Smith and Professor Kenny Coventry for all their help and support throughout the entire PhD process.

I would also like to thank all CoCo members, past and present. Special thanks to Lib Orme, who has been the best mentor anyone could ask for, Paul Agnew for his support and patience throughout, and my brother Stef for all his assistance.

I'd also like to thank Dr Natasha Kirkham and Professor Gwyneth Doherty-Sneddon for assessing this thesis. I hope they find it enjoyable.

Finally, John – thanks for everything.

Author's declaration

I declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work. I also confirm that this work fully acknowledges opinions, ideas and contributions from the work of others.

Any ethical clearance for the research presented in this thesis has been approved. Approval has been sought and granted by the School of Life Sciences Ethic Committee at the University of Northumbria.

Name:

Signature:

Date:

Chapter 1

Introduction

Die Umgangssprache ist ein Teil des menschlichen Organismus und nicht weniger kompliziert als dieser.

As part of our being, language is no less complicated than we are.

Ludwig Wittgenstein

Language is a truly intriguing human capacity. It comes naturally to humans; it emerges as a defining feature in humans regardless of their intellectual abilities and despite the lack of parental reward (Chomsky, 1965). At the same time, the grammatical structure underlying natural languages is immensely complex. This leads to a conundrum: How do infants acquire their native language with relative ease despite this complexity? There are two opposing approaches toward solving this conundrum. On one hand, the nativist view involves a specialised system within the human brain, whose one purpose is that of language acquisition. On the other hand, the empiricist approach argues for language being subserved by a set of general-purpose learning tools, which also operate across other cognitive domains, outside of language.

In this thesis, the acquisition of one particular structural complexity ubiquitous in natural languages will be investigated in detail: Non-adjacent dependencies. Non-adjacent dependencies are regularities between two elements separated by an arbitrary number of intervening elements. A frequent example found in English is that of subject-verb agreement, as in the sentence *The socks on the floor are red*. The underlined units represent the non-adjacent dependency, which in this case operates across the plural subject (*The socks*) and the verb *to be*, as these need to agree in number. If the verb did not agree with the subject, and was in the singular form, then this would produce the ungrammatical sentence **The socks on the floor is red*[1]. As the units involved in non-adjacent dependencies are interrupted within the linguistic sequence, in this case by the prepositional phrase *on the floor*, their acquisition might be expected to pose a challenge for learners. The experimental work presented in this thesis will investigate the acquisition of non-adjacent dependencies, with specific focus on whether they are learnable across domains and across different perceptual modalities, and whether additional (non-linguistic) cues facilitate their detection.

In this chapter, it will be argued that non-adjacent dependencies form an ideal test case for investigations into human language learning abilities as they represent an important structural complexity in natural languages. Further, two formalisms for

[1] Asterisks indicate ungrammaticality.

representing grammars will be introduced, one of which (Finite State Grammar) is not powerful enough to do the complexity of natural languages justice, specifically non-adjacent dependencies. Next, the nativist view on language acquisition will be contrasted against the empiricist view, and finally, language will be discussed in light of domain-specificity.

1.1 Non-adjacent dependencies as a defining feature of language

Non-adjacent dependencies are ubiquitous in natural languages. Seven examples of frequently found non-adjacencies in English are listed below. In all of them, the underlined elements are the ones forming the regularities.

(1) The socks on the floor are red.

(2) John will buy either red socks or green socks.

(3) If John won five pounds, then he would buy red socks.

(4) John is buying red socks.

(5) John is buying socks, isn't he?

(6) What did John buy _?

(7) John seems_to like socks.

Sentence (1) is the example of subject-verb agreement introduced above. The plural noun phrase *The socks* must agree in number with the verb to ensure grammaticality. In example (2) *Either* must be proceeded by the conjunction *or*. Thus, on hearing *either*, the listener expects *or* to occur at some point later on in the sentence. Example (3) shows the frequently-found non-adjacent regularity between the subordinate conjunction *If* and *then*. In (4) the non-adjacent dependency operates between the auxiliary *is* and the suffix *–ing* in order to indicate progressive aspect. The dependency in this case is separated by the stem of the main verb *buy*. The non-adjacent dependency in the tag question in (5) holds between the subject *John* and the pronoun in the tag as they need to agree in order to render the question grammatical. Examples (6) and (7) illustrate that non-adjacencies do not exclusively function between overt elements. (6) is a wh-dependency or filler-gap construction, in which the regularity holds between the underlined wh-word and a syntactically corresponding gap at the tail end of the question. Hearing the word *What* thus initiates a search for the gap. The nature of this gap is heavily disputed. In mainstream generative grammar, this gap represents the extraction site, i.e. the wh-word originated in this position and was fronted in order to form the question. If movement is not considered a possibility, then the gap nevertheless is syntactically related to the wh-word as it indicates a position which can be occupied by wh-words (as in *John bought what?*). In either case, a non-adjacent relationship holds between the two underlined elements. Crucially, one of the elements receives a null-spellout, meaning that it is neither aurally nor visually perceived (Radford, 1997). In sentence (7), *John* is analysed as the subject of the subordinate clause. As in the wh-dependency (6), under a movement analysis, the non-adjacency in this case holds between a moved element and the corresponding extraction site. The trace left at the extraction is again neither heard nor seen, receiving a null spellout.

Non-adjacent dependencies are viewed as a defining feature of natural languages based on the fact that they are a diagnostic of Phrase Structure Grammars. The following two sections will introduce two formalisms for representing grammars and explain that a Phrase Structure Grammar is better suited to describing natural languages than a Finite State Grammar, as the latter cannot adequately capture non-adjacent dependencies.

1.1.1 Finite State Grammars

It would seem plausible to assume that words are strung together, one after the other to create sentences. Finite State Grammars (FSGs) do precisely that (Chomsky, 1956). FSGs can be specified entirely by the transitions from one state to the next. The system produces one symbol for each transition. Although this is a very simple grammar using a finite number of states, it is nevertheless capable of generating an infinite number of sentences, by including loops between states. For illustration purposes, again consider the sentence (1) below:

(1) The socks on the floor are red.

A FSG would produce this sentence by creating a chain of words with a flat, linear structure, as shown in Figure 1.1.

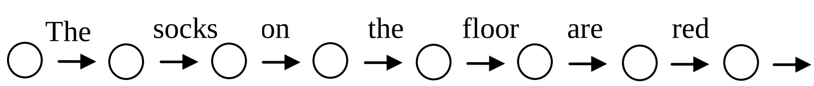


Figure 1.1: A simple FSG that can only generate the sentence *The socks on the floor are red*.

FSGs are perfectly capable of generating a more varied set of sentences as shown below (Figure 1.2).

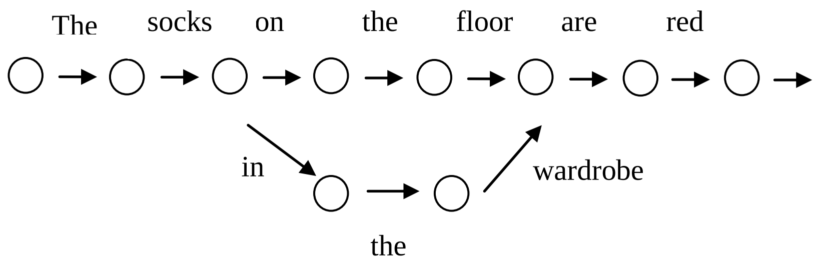


Figure 1.2: A more complex FSG that is capable of generating two separate grammatical sentences. The prepositional phrase *on the floor* can be exchanged for *in the wardrobe*.

The FSG in Figure 1.2 can capture two sentences that differ in terms of the prepositional phrase (PP): *The socks on the floor are red* and *The socks in the wardrobe are red*. Importantly, however, FSGs cannot adequately capture non-adjacent dependencies. In this sentence, the subject *The socks* is separated from the verb *are* by an intervening PP *on the floor* or *in the wardrobe*. The subject and the verb need to agree in number, and since the subject in this case is plural, the verb must also be plural. In order to generate fully-fledged human languages, the human mental grammar must be capable of capturing non-adjacencies, i.e. regularities between separated elements, efficiently while still allowing for an infinite number of new sentences. The following problem quickly arises:

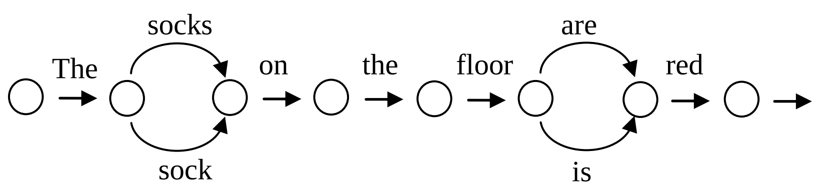


Figure 1.3: An attempt to design an efficient FSG capable of generating two grammatical sentences that merely differ in terms of their subject-verb agreement fails since it also generates ungrammatical sentences.

Although the FSG above (Fig. 1.3) can account for the singular and plural version of the same sentence, it also generates the two ungrammatical sentences **The socks on the floor is red* and **The sock on the floor are red*. In both these sentences the subject-verb agreements are being violated. The only way to overcome this problem using a FSG is to create a much larger system, such as the one shown below (Fig. 1.4):

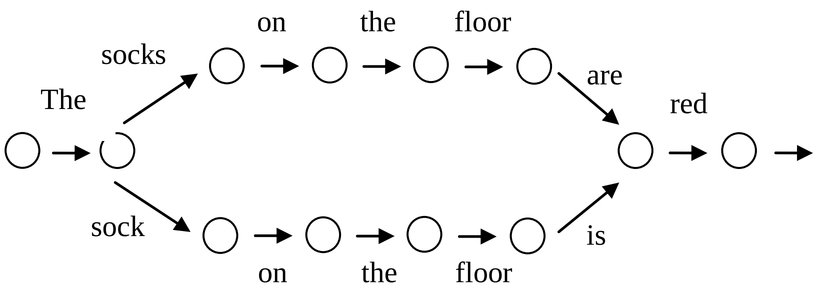


Figure 1.4: Accounting for both grammatical sentences *The socks on the floor are red* and *The sock on the floor is red* without creating grammatical violations, consequently involves FSGs becoming very large. This FSG generates exactly two sentences and already requires twelve separate transitions from one state to another.

This illustrates the problem of scaling up a FSG to capture non-adjacent dependencies. The grammar shown in Figure 1.4 is capable of generating exactly two sentences of English, and already it is fairly large. In addition to this, it is redundant and inefficient as the only true difference between the two sentences lies with the verb *to be*, and yet the entire PP has is replicated in order to avoid ungrammaticality. Thus, while FSGs are perfectly capable of generating sentences with flat structures, their inefficiency in handling complex linguistic structures that are frequently found in natural languages, such as non-adjacent dependencies, demonstrate that the mental grammar that allows humans to acquire and use natural language so effortlessly cannot be satisfactorily specified in terms of a FSG. A much more powerful system, which can describe the structural complexities found in natural languages, is therefore presented in the following section.

1.1.2 Phrase Structure Grammars

Words in natural languages are not simply strung together one after the other in order to create sentences. Rather, lexical items are grouped together to form phrases, and phrases are then organised to form sentences. As a result, sentences exhibit a hierarchical phrase structure, they follow phrase rules, and they can be generated by Phrase Structure Grammars. For illustration purposes, consider sentence (1) again:

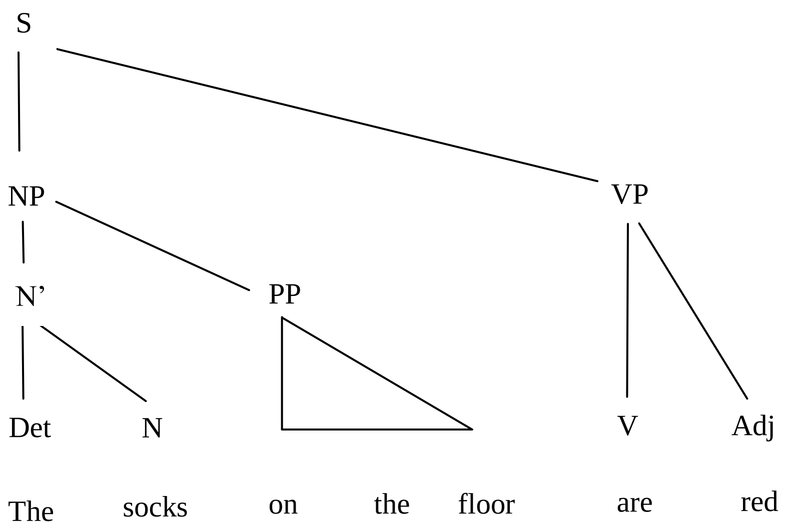


Figure 1.5: Syntactic tree structures clearly illustrate the hierarchical structure underlying natural language. The noun phrase (NP) contains the PP *on the floor.* The structure shows that non-adjacent dependency between the subject and the verb of this sentence are adjacent on a higher level of the hierarchy as the sentence S constitutes an NP and a verb phrase VP (S → NP VP).

The syntactic tree structure in Figure 1.5 exhibits the hierarchical organisation typical of language. The subject, determiner (Det) *The* and noun (N) *socks,* feeds into a larger unit, i.e. the noun phrase (NP) that contains the string *The socks on the floor*. For this reason, the non-adjacent dependency between the NP and the verb phrase (VP) *are red* are in fact adjacent further up in the hierarchy: number features are checked for higher up in the hierarchy, and therefore number agreement is in fact an adjacent relationship between the NP and VP, rather than a non-adjacent relationship between words.

The phrase structure rules in Table 1.1 complement the tree structure in Figure 1.5, and dictate that a syntactically well-formed sentence contains an NP followed by a VP, which must agree in number (indicated here by the subscript X).

Table 1.1: These phrase structure rules are capable of generating the sentence shown in the syntactic tree structure (among others) in Figure 1.5. The arrow (→) should be read as "contains". Curly brackets contain all possible elements per category. Parentheses indicate optional constituents. The agreement problem between N and Det and especially between N and V can be easily solved by specifying the number constraint. Thus, a plural NP (NP_P) can only take a plural verb (VP_P), whereas a singular NP_S can only take a singular VP_S.

$S \rightarrow NP_X\ VP_X$	
$NP_S \rightarrow Det_S\ N_S$ (PP)	$N_S \rightarrow$ {sock, floor}
$NP_P \rightarrow Det_P\ N_P$ (PP)	$N_P \rightarrow$ {socks}
$PP \rightarrow P\ NP$	$P \rightarrow$ {on}
$VP_X \rightarrow V_X\ Adj$	$V_S \rightarrow$ {is}
	$V_P \rightarrow$ {are}
$Det_S \rightarrow$ {the}	$Adj \rightarrow$ {red}
$Det_P \rightarrow$ {the}	

In addition to forming the plural version of the sentence shown in Figure 1.5, the PSG shown in Table 1.1 is powerful enough to generate additional sentences, which are all grammatically sound, for example the singular version *The sock on the floor is red*. PSGs are thus an efficient way of capturing the complexities found in natural languages, most notably non-adjacent dependencies. PSGs rather than FSG are therefore considered to be the adequate formalism to describe natural languages.

Importantly, non-adjacent dependencies are diagnostic of PSGs, and as such they make a good test case for investigations into the human language capacity. For this reason, they have been chosen for the present thesis to investigate domain-specificity in language learning.

1.2 Language Acquistion

As outlined by Bates (1994), after a phase of initial babbling, children are believed to show first signs of word comprehension between the ages of 8 – 10 months. Speech production starts at about 12 months, with the single-word stage lasting for approximately 8 months, followed by the telegraphic speech phase, during which children produce simple strings of two lexical items. By the time they reach their third birthday, children will have a fairly sophisticated knowledge of complex grammar and an extensive lexicon. Although it is argued that this linguistic knowledge is acquired not by reinforcement or extensive training but by mere exposure to the target language (Chomsky, 1965), language acquisition always takes place in a social context (Clark, 1977). It crucially depends on how adults communicate with infants, e.g. by getting the infant's attention by touching them or using a certain tone of voice, and by talking about something relevant to the infant, such as a certain toy (Clark, 1977).

Tomasello (2008) points to the importance of common ground and shared intentionality in language acquisition, also clearly placing language learning in a social context. Evidence for this, so he argues, comes from the fact that with human infants one crucial aspect of language learning, namely word learning, begins at around 12 months of age – no sooner, no later. This coincides with the emergence of shared intentionality itself, at somewhere between 9 and 12 months of age. The importance of shared intentionality, for example, was revealed in an experiment run by Tomasello, Strosberg and Akhtar (1996). In this design, an adult participants and an 18-month-old infant engaged in a finding game, in which the adult was required to express a clear intention, such as finding a *toma* within an array of entirely novel objects. The adult subsequently went through all novel objects until she found the *toma*, which she indicated by smiling and ending the search. This behaviour alone was sufficient for the infant to acquire the novel word *toma*, showing awareness of the adult's intentions.

The fact, however, still remains that human infants acquire language fairly effortlessly and within a short period of time. The following sections will focus on two alternative hypotheses why this might be the case.

1.2.1 The innate language learning device

Section 1.1.2 illustrated (some aspects of) the complex structure underlying language. However, despite this complexity, by the time children reach their third birthday, they will have a fairly sophisticated knowledge of complex grammar and an extensive lexicon (Bates, 1994). From a nativist perspective, this knowledge is acquired with the help of a very specialised mechanism. This view, along with the main supporting arguments, is presented here.

As Chomsky (1988) points out, the input from which the target language is to be acquired by each infant is impoverished in nature. This argument from the poverty of stimulus (PoS) highlights that children are neither exposed to the full richness and complexity of language, nor is every instantiation of the input grammatically sound, and yet they acquire the correct cognitive representation of the relevant grammar in its entirety. This apparent disparity between the ease of language acquisition on the one hand, and the imperfect exposure on the other has led to a number of theories regarding the cognitive tools available to human infants to assist in this task. Famously, Chomsky (1965, 1988) assumes that children tackle this challenge with an intuitive understanding of language already in place, the Language Faculty.

A well-known example Chomsky (1986) uses to construct the PoS argument deals with anaphoric reference. Anaphors are constrained by the "binding theory" (p. 8), meaning that they are structurally dependent on an antecedent.

(8) I wonder who [the men expected to see them]

(9) [the men expected to see them]

The pronoun *them* is interpreted differently in the two examples. In (8) it is referentially dependent on the preceding NP *the men*, whereas in (9) this interpretation is not possible. According to Chomsky, during language acquisition children are rarely exposed to instances of anaphoric reference as shown in examples (8) and (9), yet despite this, the correct interpretation of the pronominal anaphor is identified fairly effortlessly. In line with the PoS argument, Chomsky maintains that this linguistic knowledge (in this case: selecting the correct antecedent for the anaphora) despite the impoverished input (in this case: lack of exposure to these structural complexities) is achieved with assistance from an innate Language Faculty.

The Language Faculty is a cognitive system, which – on exposure to the primary linguistic data – allows the child to construct a theory of the target language. Once the mature linguistic state has been reached, the language learner will be able to understand and produce language based on these internal representations. Chomsky (1988) visualises this model of language acquisition in the following way (Fig. 1.6):

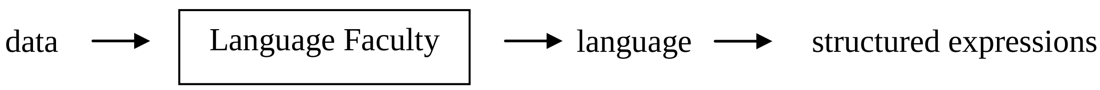


Figure 1.6: Illustration of Chomsky's language faculty. The target language (data) serves as input, the language faculty then uses this input to create a representation of the language, which in turn results in structured expressions. Taken from Chomsky, 1988.

Chomsky claims that this language-specific learning tool forms part of our biological endowment as humans. As a logical consequence, due to the fact that the Language Faculty must be able to form a language based on any human language as input, a set of universal principles is required. Chomsky calls these principles of Universal Grammar (UG). UG thus represents the state of the Language Faculty before exposure to the linguistic input, and the grammar of the target language (e.g.

English) represents the Language Faculty in its mature state. UG provides the language learner with an algorithm for forming a grammar based on experience. The principles of UG must, as a consequence of being universal, be able to constrain every single grammatical operation possible in every human language.

This innate algorithm for language acquisition must be complemented by a set of parameters that will adapt according to language-specific variation: Since natural languages vary along certain dimensions, e.g. positioning of heads, these aspects must be parameterised. As an analogy, the set of parameters can be seen as a huge switch box. Each switch can take only one of two values. Taking the head position parameter as an example, a child learning English as a native language, will specify this parameter as head-first (i.e. flick the switch towards the head-first value due to English being a head-first language), whereas a child exposed to Korean will push the switch in the opposite direction, giving it the head-last value. Thus, within the Chomskyan framework learning a language, or rather acquiring grammatical structure, is facilitated by an innate set of universal principles plus parameters which are assigned values based on environmental influences, i.e. the primary linguistic input. According to Chomsky then, the challenges children face during language acquisition are limited due to the uniquely human Language Faculty.

A logical consequence of the fact that all humans are, so claims Chomsky, equipped with the same Language Acquisition Device (LAD) is that all languages are similar. As Chomsky (1980) says:

> I have not hesitated to propose a general principle of linguistic structure on the basis of observation of a single language. (...) Assuming that the genetically determined language faculty is a common human possession, we may conclude that a principle of language is universal if we are led to postulate it as a “precondition” for the acquisition of a single language. (p.48)

The human capacity for language is frequently used as a prime example of domain-specificity. The view that the human mind is organised into specialised, domain-specific systems has been maintained from different perspectives. The following section will therefore focus on the main arguments for domain-specificity.

1.2.2 Modularity and domain-specificity

One influential accounts of domain-specificity comes from Fodor (1983), who described the cognitive organisation of the human mind as modular, arguing that individual, identifiable cognitive systems underlie human behaviour. According to this Modularity Thesis, in order to count as a module, a cognitive system must fulfil the majority or all of the following nine criteria shown in Table 1.2.

Table 1.2: Fodor's (1983) nine requirements for modularity.

1	Domain-specificity	6	Shallow output
2	Mandatory processing	7	Fixed neural architecture
3	Limited central access	8	Specific breakdown patterns
4	Speed	9	Specific pace and processing
5	Encapsulation		

Domain-specificity is thus, according to Fodor a feature of modularity, describing how modules process information. Fodorian modules operate automatically, with other cognitive systems having limited access to the internal computations of modules (criteria 2 and 3); modules are fast (criterion 4), they cannot draw on information from other systems or modules (criterion 5), and they do not provide information about the intervening levels of interaction leading to the output (criterion 6). Criteria 7 – 9 concern the architecture of modules. In accordance with Fodor, modules are represented by identifiable neural structures (criterion 7), they show patterned breakdown characteristics such as aphasia (criterion 8), and they develop in predetermined fashion with regard to pace and sequence (criterion 9). Most importantly for this thesis, however, is the first requirement: Domain-specificity. Fodor reasons that modules are domain-specific and are thus "special-purpose" (p.51) mechanisms, dealing only with stimuli of a certain type. He uses the human capacity for language as an example to strengthen the case for domain-specificity, arguing that the linguistic universals found across languages reflect language-specific learning biases. This clearly demonstrates the intrinsic relationship between innateness and domain-specificity that many people see. The human ability to acquire language so effortlessly is therefore, in line with Fodor, due to a domain-specific, isolated learning mechanism aimed at processing linguistic stimuli only. Although Chomsky's view does not explicitly involve encapsulation of the LAD, Fodor and Chomsky's approaches are nevertheless similar.

Domain-specificity is then, in the Fodorian framework, a feature of modularity: Modularity includes domain-specificity. Definitions of modularity, and to what extent modules fulfil the requirements Fodor (1983) stipulated are heavily discussed and vary hugely. Many views of modularity or the organisation of the human mind support a much weaker version of the Fodorian Modularity (Carey & Spelke, 1994; Carruthers, 2006; Cosmides & Tooby, 2006).

Similar to Pinker and Bloom (1990), evolutionary psychologists such as Carey and Spelke (1994) and Cosmides and Tooby (1994, 2006) argue for the human mind as an adaptation to the environment, evolved under selective pressures much like any other organ. Carey and Spelke (1994) as well as Spelke and Kinzler (2007) assume a suite of core principles, and domain-specific knowledge to be arranged around these principles. For the human capacity of language, this core set is instantiated as the principles of UG. In their view then, the human mind is not made up of exclusively encapsulated modules in the Fodorian sense, and neither is it organised in a completely general-purpose fashion. They do, however, assume specialisation of cognitive systems to some extent, and explicitly so for the human language capacity.

Tooby and Cosmides (1992) view the mind similarly encapsulated. In line with Fodor, who views encapsulation as a defining feature of modularity, Tooby and Cosmides consider encapsulation to be a key argument in ruling out domain-generality. According to Tooby and Cosmides, cognitive mechanisms must be informationally encapsulated in order for them to function efficiently. If this requirement was not met, perceptual mechanisms would have to check with an extraordinarily large amount of cognitive resources that might all be relevant to the processing of an incoming signal. This would therefore not allow for time-efficient computations. In their evolutionary-based view, specific adaptive problems can only be confronted by specialised mechanisms since “generality can be achieved only by sacrificing effectiveness” (Cosmides & Tooby, 1994, p.89), and domain-specific mechanisms “systematically outperform (...) more general mechanisms” (p.89).

A weakened version of Fodorian modules are also claimed by Carruthers (2006). Carruthers argues for massive modularity, meaning that the entire mind is made up of distinct components, each of which is responsible for a very specific task. However, Carruthers' modules are not exclusively domain-specific, do not have shallow outputs and need not be very fast, unlike Fodor's, as they include conceptual beliefs and thoughts. Carruthers' modules are therefore weaker than the Fodorian modules, and thus do not fulfil the nine properties listed above. However, considering the sheer amount and variety of stimuli people are faced with on a day-to-day basis, Carruthers finds it hard to assume a domain-general mechanism would be able to process them. He thus finds it much more likely that individual components display a certain amount of task-specificity.

The Fodorian Modularity Thesis has also been explicitly challenged. By criticising each of Fodor's nine principles in turn, Prinz (2006) concludes that modularity is an inadequate view of the human mind, particularly modularity in the Fodorian sense. Rather, the mind should be viewed as a "network of interconnected systems and subsystems" (p.33). This shows that his main criticisms regard encapsulation and domain-specificity. Prinz argues against domain-specificity, although he acknowledges that supposed modules might well have domain-specific components. In language, for example, highly localised dedication to tasks, such as conjugating irregular verbs, might well fall within alleged modules, but this by no means implies language as a whole is a domain-specific module in the strict sense. Prinz is, however, not as adverse to massive modularity, which is by its very definition a less encapsulated perspective on modularity. Thus, Prinz's approach to the non-modular organisation of the mind acknowledges that associated systems and subsystems will nevertheless allow for high task-specificity.

There are thus a number of different views of modularity and of domain-specificity, and all the in-principle approaches discussed thus far assume that the human mind may be to an extent specialised, and a key argument for this seems to be the fact that domain-general mechanisms are assumed to be inefficient (Cosmides & Tooby, 1994). Applying this to the human language abilities, this would support Chomsky's view of a domain-specific device, whose sole purpose is language acquisition. There is, however, an opposing view to the innate, language-specific learning device, which will be presented in the following sections.

1.3 Is there a need for an innate language learning device?

Assuming the existence of an innate Language Faculty bridges the gap between the limited input and the linguistic capacity of humans in the mature state. However, innateness and the idea of UG and LAD as put forward by Chomsky have been challenged from a number of different lines of research. Specifically, it has been questioned whether UG could have evolved in the first place (Christiansen & Chater, 2008) whether all languages in the world are in fact similar (Evans & Levinson, 2009), and whether the PoS is indeed a good enough argument for LAD (Kirby, Cornish & Smith, 2008; Pullum & Scholz, 2002; Smith, Brighton & Kirby, 2003; Smith, Kirby & Brighton, 2003). However, if the idea of an innate language learning tool in the Chomskyan sense is dismissed, as is suggested here, then the linguistic knowledge humans demonstrate must be achieved with assistance from efficient general-purpose learning tools (Elman, Bates, Johnson, Karmiloff-Smith, Parisi & Plunkett, 1996).

Recent years have seen an increase in artificial language learning (ALL) and artificial grammar learning (AGL) experiments, which have identified a number of powerful cognitive tools that assist humans in language learning[2]. The main issue with examining language and language abilities in humans is that language acquisition itself starts at a very early age, possibly even in utero (e.g. DeCasper & Spence, 1986; Busnel, Granier-Deferre & Lecanuet, 1992). Thus, when taking part in experiments, even infants will already be equipped with prior linguistic knowledge to some extent. For this reason, most studies of this nature involve either an ALL or AGL paradigm, which can be designed by the researcher to replicate the linguistic structure that is to be examined. ALL/ AGL experiments are therefore an accepted method for investigating human linguistic abilities for two main reasons. Firstly, they allow for experimental control as the language/grammar can be designed specifically to target the phenomenon of interest, and secondly, it allows the experimenter to rule out the possibility of participants' previous exposure to the language/ grammar itself. ALL and AGL experiments usually involve a (more or less) implicit training phase, during which participants are exposed to instantiations of the target language/ grammar. This training is usually followed by a testing phase, in which the acquired knowledge is assessed.

With the assistance of ALL and AGL paradigms, a number of domain-general learning mechanisms have been exposed, which humans can employ in language acquisition as well as in other learning tasks, and which therefore represent a true alternative to an innate, language-specific acquisition device. It has been shown that infants are, from a very early age, sensitive to the statistical structure

[2] With regard to terminology, there is no real consensus in the research area on what experimental design constitutes an ALL and what paradigm constitutes an AGL experiment, and frequently both expressions are used synonymously. For the purposes for this thesis, however, experiments investigating language learning more broadly or how meaning is mapped onto words will be referred to as ALL experiments. AGL experiments are, in the present context, experiments investigating the acquisition or learnability of an underlying grammar (e.g. a FSG or a PSG), regardless of whether the grammar is realised as linguistic or as non-linguistic materials.

underlying a sequential input. Kirkham, Slemmer and Johnson (2002) demonstrated that 2 month-old infants are sensitive to shape-to-shape conditional probabilities. By the age of 8 months, infants can reliably identify words within the speech stream, relying solely on the co-occurrence of neighbouring syllables (Saffran, Aslin & Newport, 1996), and by 12 months, they are capable of extracting a FSG after minimal exposure (Gómez & Gerken, 1999). Thus, by the time human infants reach the age of 1, they are adept at harnessing element-to-element transitions in order to make sense of sequential stimuli. However, as explicated in 1.1.2, natural languages exhibit hierarchical structure. Sensitivity to structural regularities, which are based on syntactic hierarchy, has been identified at a later stage in development. Gómez and Maye (2005) showed that, in an AGL paradigm, 15 month-old infants can successfully track non-adjacent dependencies between artificial words, whereas 12 month-olds cannot. This reflects the addition cognitive challenge involved in tracking non-adjacent dependencies. These mechanisms will now be discussed in detail, with specific focus on the strategies used in order to identify words within the fluent speech stream (Section 1.3.1), to learn simple FSGs (Section 1.3.2), to abstract rules and apply them to novel instances of language (Section 1.3.3), and to learn non-adjacent dependencies (Section 1.3.4).

1.3.1 Statistics and segmentation

In spoken language, there are no obvious pauses between individual words (Saffran, Aslin and Newport, 1996). For this reason, one of the first challenges an infant faces in language acquisition is to identify individual words by breaking down the fluent speech stream into meaningful units. A number of techniques have been suggested, which might assist infants in this task. Jusczyk (1999), for example, suggests that infants can pick up on allophonic cues or prosodic patterns of the target language to detect word boundaries. Jusczyk, Houston & Newsome (1999) found that infants as young as 7.5 months are sensitive to the strong-weak stress pattern found in English and are reliably able to use this regularity to segment bisyllabic words. Similarly, Mattys, Jusczyk, Luce & Morgan (1999) found that prosodic as well as phonotactic cues are available to 9-month-olds for the detection of word boundaries.

In addition to showing sensitivity to these purely linguistic cues, infants are also sensitive to the statistical structure underlying the speech stream. Moreover, they are capable of harnessing this structure and using it to segment the fluent speech stream into words. This was demonstrated in a classic experiment conducted by Saffran, Aslin and Newport (1996). In their experiment, 8-month old infants were exposed to a 2 minute, aurally presented artificial language, which consisted of a total of four artificial words (*bidaku*, *golabu*, *tupiro* and *padoti*), strung together in random order. Importantly, the speech stream lacked prosodic cues or pauses to indicate word boundaries. Thus, the statistical structure of the speech stream was the only indicator of where one word ended and another began. The infants were aware of the syllable-to-syllable transitional probabilities of the briefly-presented language, and could subsequently differentiate between high- and low- probability transitions. This was assessed by measuring the infants' listening times during the test trials. Specifically, the infants were capable of harnessing the fact that the transitional probabilities within words, such as *bi* → *da* were p = 1.0, as these are neighbouring syllables in the word *bidaku*. By contrast, transitional probabilities spanning word boundaries were lower. For example, the two words *bidaku* and *golabu* strung together form the sequence *bidakugolabu*. Transitional probabilities between the adjacent syllables *ku* → *go* was p = .33, as the word *bidaku* (and therefore the syllable *ku*) could have been proceeded by any one of three words (and therefore any one of three syllables). On test, the infants were capable of computing the transitional probabilities between adjacent speech sounds to differentiate between words and non-words (i.e. an entirely new combination of three syllables that were encountered during the training, e.g. *tilado*). Impressively, infants were also able to discriminate between and words and part-words. Part-words were three-syllable sequences, which were produced by proceeding the final syllable of one word with the two initial syllables of another word. Part-words, unlike non-words, occurred during the training, which demonstrates that the mechanism underlying this ability must be very sensitive. This type of statistical learning has also been demonstrated in adults (Saffran, Newport & Aslin, 1996).

Tracking transitional probabilities is not a language-specific mechanism. Saffran, Johnson, Aslin and Newport (1999) conducted a series of experiments to test whether this mechanism is also accessible in the non-linguistic domain. To do this, the sequences of the artificial language from their previous studies were translated into musical tones: each artificial word corresponded to a 'tone word' in that each syllable from the AL was matched with a specific tone. Thus, the syllable *bu*, for example, was substituted with the tone D. Their results indicate that adults as well as 8-month old infants can reliably extract the tone-to-tone transitional probabilities to segment the tone stream.

Furthermore, Kirkham, Slemmer and Johnson (2002) demonstrated that the same results obtain in the visual modality. In their experiment, 2-, 5- and 8-month old infants were familiarised with sequences made up of a total of six shapes (turquoise square, blue cross, yellow circle, pink diamond, green triangle and red octagon). Each stimulus therefore differed from the next along the two dimensions of colour and shape. The shapes were presented individually on a computer screen for 1 second, which replicates the sequential nature of the aurally presented language used by Saffran, Aslin & Newport (1996). Importantly, the sequences were structured based on the transitional probabilities of specific shape-pairs. The square was reliably followed by the cross with a transitional probability of $p = 1.0$, and the same was true for the shape pairs circle-diamond and triangle-octagon. There was thus a total of three shape-pairs, and therefore transitional probabilities spanning two pairs were at $p = .33$. On test, the infants viewed sequences of pairs that followed the pattern from the training, as well as sequences in which the shapes were randomly ordered. The infants showed a novelty effect with looking significantly longer at the novel sequences. This was true for each age group. Kirkham et al. therefore convincingly demonstrated that the capacity to calculate transitional probabilities between adjacent elements also operates in the non-linguistic domain. Similar findings by Kirkham, Slemmer, Richardson and Johnson (2007), and Fiser and Aslin (2002b) demonstrate infants' sensitivity towards adjacent relationships between elements in multiple locations and in multi-element scenes, respectively.

Conducting an ERP experiment, Abla and Okanoya (2009) replicated Saffran, Johnson, Aslin and Newport's (1999) experiment by replacing each tone with a shape. Thus, each 'shape word' in this experiment included three shapes, and six shape words were strung together in random order to form a sequence for the training phase. The test involved participants being presented with two three-shape strings: a shape word and a shape non-word. Participants were then required to decide which one was familiar to them. The results showed that participants here were also capable of successfully segmenting the shape stream based on unit-to-unit transitional probabilities.

Research into the statistical learning and grammar learning abilities of animals can provide additional insights into the issue of domain-specificity with regard to language learning. Natural language in all its complexity is only found in humans (Hauser, Chomsky & Fitch, 2002[3]). As a consequence, any sequencing capacities or learning abilities identified in non-human animals cannot be unique to human language, and thus not language-specific in this sense. For this reason, research on the cognitive abilities of non-human primates (Hauser, Newport & Aslin, 2001) and birds (Takashi, Yamada & Okanoya, 2010; Ohms, Gill, Van Heijningen, Beckers & de Cate, 2010) can offer insights into the nature of human learning mechanisms. For the present context, research on non-human primates, specifically on stream segmentation is particularly relevant.

Hauser, Newport and Aslin (2001) investigated speech stream segmentation in cotton-top tamarins (*Sanguinus oedipus*) by employing the materials and a similar experimental set-up as used by Saffran, Aslin and Newport (1996). Their findings show that tamarins are capable of harnessing the underlying statistical structure of a continuous speech stream to discriminate words from non-words and part-words, detecting word boundaries based on transitional probabilities much like human infants and adults. This suggests that tracking transitional probabilities for stream segmentation is not a mechanism unique to human language.

[3] Due to the allegations regarding M. Hauser's scientific misconduct, the retracted 2002 paper published in *Cognition* is not discussed in the present thesis. However, other research by Hauser will be considered where relevant.

However, it may be argued that statistical learning (by domain-general learning mechanisms) can only be employed for relatively low-level processes, and that higher-level processes required for language learning might nevertheless be supported by domain-specific learning tools. The following sections will address this, and demonstrate that this is not the case.

1.3.2 Learning FSGs

Language learning goes beyond identifying words within the speech stream and mapping meaning on to words. Thus, in order for humans to acquire fully-fledged natural languages, they must be capable of more than tracking statistics – they must also be able to learn and parse complex sequential structure.

Using the headturn preference procedure (HPP), Gómez and Gerken (1999) showed that 1-year-old infants are capable of learning sequences generated by a FSG. The HPP (Kemler Nelson, Jusczyk, Mandel, Myers, Turk & Gerken, 1995) is a reliable tool to investigate infants' perception of aurally presented stimuli. Usually, it requires the infant to be seated on the caregiver's lap in a darkened booth. The infant's attention is initially drawn towards a flashing light opposite the infant. One additional light is placed at either side of the booth. Once the centre light stops flashing, one of the side lights starts flashing to indicate that the auditory stimulus will be presented shortly. When the infant directs his/her attention to the side light, the stimulus starts playing. The infants' looking times to each of the stimuli are then recorded and subsequently analysed. Gómez and Gerken first acquainted infants with instantiations of the FSG shown in Figure 1.7.

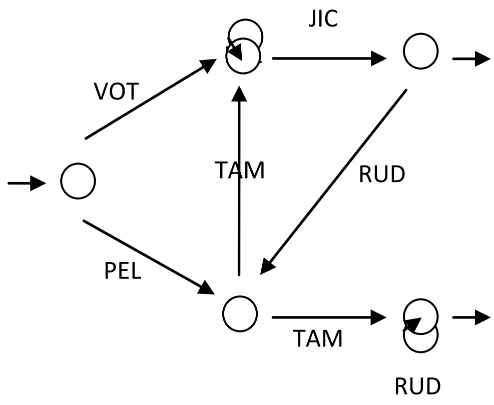


Figure 1.7: FSG (G1) used by Gómez & Gerken. G2 (also an FSG) used for the experiment is not depicted here.

In subsequent tests, they showed that infants were reliably able to discriminate grammatical from ungrammatical sequences, regardless of whether the grammatical violation involved sequential endpoints or sequence-internal transitions from one word to the next. Moreover, Gómez and Gerken demonstrated that infants were able to abstract the FSG and apply that knowledge to instantiations of the same grammar with new vocabulary.

This is an important step in language acquisition as language cannot be merely memorised, but rather infants must abstract the rules governing legal sequences and apply them to novel instances of the target language. Abstract rule learning is thus a crucial aspect in language acquisition and will be discussed in more detail in section 1.3.3.

FSGs have also been shown to be learnable in the non-linguistic domain, with the grammar instantiated as just black (Conway & Christiansen, 2005) or coloured squares (Conway & Christiansen, 2006; Conway & Christiansen, 2009), tones (Conway & Christiansen, 2005; Conway & Christiansen, 2006; Conway &

Christiansen, 2009) and vibrotactile stimuli (Conway & Christiansen, 2005). Thus, learning a FSG grammar is not merely possible in the linguistic domain.

Moreover, the ability to learn sequences generated by a FSG has been demonstrated in non-human primates. Fitch and Hauser (2004) investigated tamarins' ability to discriminate between grammatical and ungrammatical strings generated by either a FSG or a more complex PSG. In their experiment, both humans and tamarins were trained and tested on the same aurally presented materials. Subjects (tamarins and adult humans) were trained on either strings following a FSG or on strings generated by a PSG. In the FSG condition, the strings followed an $(AB)^n$ grammar, and for the PSG grammar, the sequences took an A^nB^n form. Each category (A and B) were instantiated as a CV syllables. In the FSG condition, therefore, each A syllable could be followed by any B syllable, and this could be repeated a number of times. The A^nB^n grammar (PSG) generated centre-embedded sequences as n A syllables had to be followed by precisely n B syllables. This created sequences of the form $A_1A_2A_3B_3B_2B_1$, in which A_3B_3 is embedded within A_2 - B_2, forming $A_2[A_3B_3]B_2$, which in turn is embedded in the outmost frame, forming $A_1[A_2[A_3B_3]B_2]B_1$. All syllables in category A were spoken by a female voice, whereas all syllables from category B were spoken by a male voice. Fitch and Hauser found that their human participants were reliably able to discriminate between grammatical and ungrammatical sequences on test in both the FSG and PSG condition. The tamarins, by contrast, successfully learnt the FSG, but failed to discriminate between grammatical sequences and strings containing violations when trained and tested on the PSG. Fitch and Hauser therefore argue that extracting a FSG is not a purely linguistic, and possibly not even a purely human, capacity.

1.3.3. Abstract rule learning

Marcus, Vijayan, Bandi Rao and Vishton (1999) investigated abstract rule learning in the linguistic domain. In their experiments, 7-month-olds were tested using the HPP. Here, they were trained and tested on a simple ABA/ABB language, and each variable was instantiated as a CV ‘word’. ABA strings thus took the form of *ga ti ga* or *li na li*, whereas ABB strings were instantiated as *ga ti ti* and *li na na*. The ABA grammar provided the ungrammatical test items for the ABB-trained infants and vice versa. Crucially, however, infants were tested on an entirely novel set of CV words, and were thus not able to use statistical learning mechanisms, such as transitional probabilities between specific elements, to identify grammatical sequences. The results indicate that infants can reliably discriminate between ABA and ABB patterns, and must therefore have internalised the abstract patterns embodied in their training data.

Abstract rule learning across domains to date has rendered a mixed pattern of results. Marcus, Fernandes and Johnson (2007) showed that abstract rule learning, in the auditory modality, is not as easily accomplished in other domains, such as varying timbres, animal sounds and musical tones, unless infants are familiarised with linguistic instantiations of the patterns first. In the visual modality, however, Saffran, Pollak, Seibel and Shkolnik (2007) showed that infants at 7 months of age are capable abstracting algebraic rules in non-linguistic sequences. In Saffran et al.’s series of experiments, the non-linguistic materials were pictures of dogs and cats, and therefore familiar stimuli to the infants. In Marcus, Fernandes and Johnson’s study (2007), by contrast, the infants were presumably not familiar with the non-linguistic materials prior to taking part, which might have impacted on performance in as much as that they might simply not have recognised the repeated items as instantiations of identical units (the issue of stimulus familiarity will be addressed in detail in Chapter 6). Thus, abstract rule learning, while not being an entirely domain-specific ability, seems to subject to specific constraints, such as stimulus familiarity (Saffran, Pollak, Seibel & Shkolnik, 2007).

Johnson, Fernandes, Frank, Kirkham, Marcus, Rabagliati and Slemmer's (2009) findings suggest a further difficulty in abstract rule learning in non-linguistic domains. In their experiments, both 8-month-old and 11-month-old infants find it hard to encode a non-adjacent repetition pattern, ABA, using coloured shapes as non-linguistic materials. Although ABA patterns have been shown to be acquired in the linguistic domain (Marcus et al., 1999), Johnson et al. point out that non-adjacencies are generally considered to be a challenge for learners. This issue will therefore be dealt with in the following section.

1.3.4 The acquisition of non-adjacent dependencies

Johnson, Fernandes, Frank, Kirkham, Marcus, Rabagliati and Slemmer's (2009) experiments show that non-adjacent repetition patterns are not readily acquired in the non-linguistic domain. However, it is not the case that they are not acquired at all. They found that 8 month-olds trained on an ABB grammar were reliably able to discriminate an ABB from an ABA pattern on test. 11-month-olds trained on an AAB grammar, correctly discriminated AAB from ABA sequences, which 8-month-olds failed to do. Crucially, when trained on an ABA grammar, neither age group was able to distinguish between ABA and ABB sequences during the testing phase. These findings indicate that infants are sensitive to repetition structure. Specifically, infants are capable of distinguishing adjacent repetitions from non-adjacent repetitions, with 11-month-olds showing better awareness of the positioning of the repetition. When trained on non-adjacent repetitions, however, they fail to learn the pattern. As pointed out above, language learning involves more than tracking repetitions of identical elements. Crucially, non-adjacent dependencies (rather than repetitions) are a diagnostic of PSGs.

The acquisition of non-adjacent dependencies has previously been found to be challenging. Whereas infants are competent at speech segmentation at the age of 8-months (Saffran, Aslin & Newport, 1996), and capable of learning a FSG at 12-months (Gómez & Gerken, 1999), they do not seem to be capable of learning non-adjacent dependencies before they are approximately 15 months of age (Gómez & Maye, 2005[4]). Unlike in speech segmentation or in acquiring an FSG, tracking non-adjacencies requires the learner to detect regularities between elements separated by an arbitrary number of intervening items, which is an additional cognitive strain and may therefore explain the relatively late acquisition.

The acquisition of non-adjacent dependencies also seems to be subject to specific constraints. Santelmann and Jusczyk (1998) explored the ability of 18-month old infants to detect non-adjacent regularities between the verb *is* and progressive *–ing* across a varied set of intervening elements. Employing the HPP, they tested infants' preference for grammatical sentences, such as *The archaeologist <u>is</u> digg<u>ing</u> for treasures* (note that the underlined elements indicate the non-adjacent dependency of interest) over sentences such as **The archaeologist <u>can</u> digg<u>ing</u> for treasures*, in which a violation of the non-adjacency creates an ungrammatical

[4] See below for a description of Gómez's (2002) method, which is the same as used by Gómez and Maye (2005).

sentence. A crucial manipulation in their design involved the number of intervening elements. In the example above, the dependency is separated by the verb stem. However, Santelmann and Jusczyk additionally tested infants' ability to extract the non-adjacencies across a number of additional syllables, as in *The archaeologist is/ *can energetically digging for treasures*. They found that infants only track the dependency when the elements in question are separated by one syllable only. They argue that limited working memory capacities may affect the tracking of regularities over longer distances. Thus, the distance between relevant units of analysis, and limited processing capacity delay the acquisition of non-adjacent dependencies.

The correct identification of non-adjacent dependencies has also been found to be challenging for adults. In 2002, Peña, Bonatti, Nespor and Mehler found that adult speakers of French can only extract non-adjacent dependencies in linguistic sequences when pause cues assist in their detection. Peña et al. designed an 'AXC language', in which all variables were realised as CV syllables. Each syllable from category A reliably predicted a specific CV syllable from category C, with one intervening element X, thus creating one trisyllabic AXC word. For example, A_i was realised as [pu], and C_i as [ki]. The *pu_ki* frame could take any one of three syllables from category X, which produced the three distinct AXC words [puliki], [puRaki] and [pufoki]. There were two additional frames, namely $A_j\_C_j$ and $A_k\_C_k$, instantiated as *be_ga* and *ta_du*. Each of these frames could also take any of the 3 X syllables, generating a total of nine artificial trisyllabic words. Participants were trained on aurally presented sequences involving all nine words. On test, participants were required to distinguish between words and part-words of the form C_kA_iX or XC_iA_j. Although participants correctly discriminated between words and part-words when tested on the training items, they failed to do so when the words in the testing phase involved a novel X element. However, the novel X elements Peña et al. used were not entirely novel but rather, they were taken from the set of syllables from category A or C. To clarify, the "new" X syllables were realised as [be] or [ta], for example. Crucially, Peña et al. did not test for the acquisition of non-adjacent dependencies by including violations of the AXC form, such as A_iXC_j, in the testing phase. Rather, they tested for speech stream segmentation based on the transitional probabilities between non-adjacent syllables. They conclude that their participants failed to extract the underlying non-adjacent rule and apply it to words with a novel surface form. However, the inclusion of subliminal 25-ms pauses between individual words during training resulted in successful discrimination between words and part-words, even when 'new' X syllables were used. This series of experiments demonstrates the difficulty involved in processing non-adjacent dependencies as opposed to the relative easy with which adjacent regularities are computed.

In a similar experiment involving stream segmentation, Newport and Aslin (2004) found that non-adjacent dependencies between similar elements are easier to detect. Their experiment involved three different types of AL. In the Syllable Language, the non-adjacent dependencies followed an AXB form and operated between two CV syllables across one intervening CV syllable. In the Consonant Language, the non-adjacencies held across consonants, skipping vowels, as in the artificial words *dokube* and *pogute* (the underlined elements represent the non-adjacent dependencies). In the Vowel Language, the vowels formed the non-adjacent dependencies and they were interleaved by consonants, as in *pagute* and *pogitae* (again, the non-adjacencies are underlined). They found that their participants were only capable of acquiring the non-adjacencies in the Consonant and Vowel

Language, which they interpret in line with the Gestalt Principle of Similarity. For this reason, this will be explicated in detail in Chapter 2.

Considering these findings, it would seem striking that humans so successfully acquire language, in which non-adjacencies are ubiquitous. However, in Peña et al.'s design, their category X merely involved three CV syllables. In natural languages, non-adjacent dependencies occur over a much more varied set of intervening items. For example, the dependency between the two morphemes *is* and progressive –*ing* allows for any verb stem to be inserted in the intervening position, creating an extensive list of possible phrases of the form *is singing, is thinking, is dancing*, and so on.

This high variability in the intervening position was explored by Gómez (2002). She showed that 18-month-old human infant and adult participants are capable of learning non-adjacent regularities in an AGL paradigm, provided that adjacent regularities are too unreliable to allow rule extraction. Her artificial language (AL) involved three-word sequences. The first word of each sequence reliably predicted the third word, whereas the intervening, second word varied freely. Thus, she investigated the acquisition of non-adjacent dependencies on the syntax-level. Table 1.3 gives the grammar she used for both adult and infant participants.

For adult participants, there were three different non-adjacent dependencies, *pel* X *rud*, *vot* X *jic* and *dak* X *tood*. The crucial manipulation involved the set sizes the intervening X elements were taken from: Set size X_2 included the two artificial words *wadim* and *kicey*. Thus, adult participants trained and tested on set size X_2, were exposed to the sequences *pel wadim rud, pel kicey rud, vot wadim jic, vot kicey jic, dak wadim tood* and *dak wadim tood*, with the non-adjacencies always holding between a specific word in initial position and a specific word in final position. The largest pool was made up of 24 different words (see Table 1.3), which generated a total of 3 (dependencies) * 24 X words = 72 three-word sequences.

Table 1.3: Gómez's grammar (2002) for adult and infant participants, with lexical instantiations for categories a – d and X.

adult participant		infant participants	
Language 1	Language 2	Language 1	Language 2
S_1 → aXd S_2 → bXe S_3 → cXf	S_1 → aXe S_2 → bXf S_3 → cXd	S_1 → aXd S_2 → bXe	S_1 → aXe S_2 → bXd
a → pel b → vot c → dak		d → rud e → jic f → tood	
X_2 → {*wadim, kicey*} X_6 → {*wadim, kicey, puser, fengle, coomo, loga*} X_{12} → {*wadim, kicey, puser, fengle, coomo, loga, gople, tapsu, hiftam, deecha, vamey, skiger*} X_{24} → {*wadim, kicey, puser, fengle, coomo, loga, gople, tapsu, hiftam, deecha, vamey, skiger, benez, gensim, feenam, laeljeen, chila, roosa, plizet, balip, malsig, suleb, nilbo, wiffle*}			

In Gómez's experiment, participants were initially trained on the aurally presented AL shown in Table 1.3. Adults were subsequently asked to complete a testing phase, in which they were required to make grammaticality judgments for individual strings: Ungrammatical strings for participants trained on Language1 were taken from Language2, and vice versa. The grammatical violations within the ungrammatical strings were therefore quite subtle and merely involved replacing the final artificial word from one dependency with the final word from another. The test items for infant participants were generated in the same way, and they were tested by monitoring their looking behaviour using the HPP. The results show that the more variable the middle element (i.e. the larger the set size from which the X items were drawn), the more likely participants were to learn the non-adjacent dependencies.

Gómez therefore suggests that people seek invariant structure in the input. High variability of the middle element disrupts reliability of transitional probabilities between adjacent words. Participants are then, with increasing set size, more inclined to reject adjacent regularities in favour of more constant ones, in this case non-adjacent dependencies. Thus, the strategy employed by participants in this experiment is akin to the reduction of uncertainty hypothesis (Gibson, 1991): Learners will exploit the most reliable information available to them in the input.

There have been further investigations into the effect of variability on the detection of non-adjacent dependencies (Onnis, Christiansen, Chater & Gómez 2003; Onnis, Monaghan, Christiansen & Chater, 2004). Interestingly, it has been shown that reversing variability of the middle element to the non-adjacent frame elements also result in successful acquisition (Onnis, Monaghan, Christiansen & Chater, 2004). Specifically, Onnis et al. found that adults extracted the non-adjacencies shown in Table 1.3 (p.35) for a set size of 1 for the intervening X elements. Thus, in this design, there were three non-adjacent dependencies that could occur with only one middle item X. This extension of Gómez's (2002) experiment shows that as long as adjacent regularities are sufficiently unreliable, people's attention will be directed toward other, more informative regularities. Under these circumstances, non-adjacent dependencies in the linguistic domain can thus be acquired.

However, there are some shortcomings in the materials employed by Gómez (2002). Firstly, the testing phase involved grammatical sequences that were identical to the ones from the training and on ungrammatical sequences containing violations regarding the final element. Importantly, however, being able to make generalisation to new instantiations based on an extracted grammar is a necessity in language learning. Yet, Gómez did not test this ability as she did not include novel X items in the testing phase. This will be rectified in the following chapter. Secondly, in Gómez's AL, the words participating in the non-adjacencies are shorter than the words in category X. These cues can be interpreted in terms of the Gestalt Principle of Similarity, which has also been suggested as a requirement for the detection of non-adjacent dependencies (see Chapter 2, and also Newport & Aslin, 2004).

Domain-specificity with regard to the learnability of non-adjacencies has been addressed by Creel, Newport & Aslin (2004) as well as Gebhart, Newport and Aslin (2009), who investigated the acquisition of non-adjacent dependencies between non-linguistic sounds. Creel et al. explored adult participants' ability to segment a stream of musical tones based on transitional probabilities between non-adjacent sounds, and found that the regularities were indeed detected if the relevant units were similar to each other in either pitch or timbre. Gebhart et al. used non-musical sounds, and also found that non-adjacencies were detected. These experiments will be the focus and discussed in detail in Chapter 5.

Moreover, the findings by Fitch and Hauser (2004) with regard to cotton-top tamarins' inability of extracting embedded structures of the form A^nB^n in their AGL task is of crucial importance here, as this can be taken to show that these structures are simply too complex for tamarins to learn. In the $A_1A_2A_3B_3B_2B_1$ sequences, the outmost frame $A_1 – B_1$ as well as the first embedded pair $A_2 – B_2$ form non-adjacent dependencies, and are the equivalent to the English sentence *The mouse* [*that the cat* [*that the dog chased*] *ate*] *was tiny*. Importantly, tamarins failed to extract these regularities. This may thus seem to leave open the possibility of non-adjacencies being unique to humans, and therefore unique to human languages (Friederici, 2004).

Further, Friederici, Bahlmann, Heim, Schubotz and Anwander (2006) conducted a functional MRI experiment on the two grammars employed by Fitch and Hauser to explore which brain regions were engaged when processing a FSG ($(AB)^n$) as opposed to a PSG (A^nB^n) in human adults. They found that the frontal operculum was activated when completing the FSG task and that Broca's area was responsible when processing the PSG. Most importantly, the frontal operculum constitutes a phylogenetically older part of the human brain. It is thus interesting that tamarins and humans are capable of processing sequences generated by a FSG, which can be defined entirely by transitional probabilities between adjacent elements, and that precisely these computations are subserved by a phylogenetically older part of the human neuroanatomy. Hierarchically structured sequences, by contrast, which – in Fitch and Hauser's experiment – are not learnable by tamarins, are processed in a comparatively younger part of the human brain. However, Friederici et al. very carefully interpret their results merely with regard to the neuroanatomical differentiation between the two systems. They also point out that it is too early to make hasty judgements with regard to which aspects are unique to humans based solely on these results.

In light of these findings it is particularly interesting that non-adjacent dependencies have been found to be acquired by cotton-top tamarins under specific conditions. In an experiment by Newport, Hauser, Spaepen and Aslin (2004), the tamarins were familiarised with an aurally presented continuous stream of grammatical sequences based on non-adjacent regularities, and tested on grammatical and ungrammatical strings. They used the same three different types of an AL as Newport and Aslin (2004), i.e. a Syllable Language, a Consonant Language and a Vowel Language. Newport Hauser, Spaepen and Aslin showed that the tamarins reliably differentiated between words and part-words when familiarised and tested on the Vowel Language and Syllable Language. Newport et al. argue that the reason for the tamarins failing to encode the regularities in the Consonant Language might be related to the fact that consonants are not as acoustically prominent as vowels and syllables, especially so for tamarins. Consonants might therefore not be sufficiently salient for tamarins to compute the relevant regularities. The most interesting finding, however, is that tamarins can segment a continuous stream based on non-adjacent syllables whereas humans cannot[5] (Newport & Aslin, 2004). Newport et al.'s most convincing explanation for this result is that tamarins may have analysed the stimulus materials holistically, based on salient features, and thus detected the non-adjacent regularities between syllables and vowels. In the present context, however, Newport et al.'s experiment is important with regard to domain-specificity: cotton-top tamarins are spontaneously capable of segmenting a continuous stream of speech based on transitional probabilities between non-adjacent elements. This result suggests that the mechanism underpinning this ability is therefore not restricted to human language.

1.4 Re-visiting domain-specificity

The experimental evidence for powerful domain-general learning tools is overwhelming and hugely reduces the need to hypothesise an innate LAD. However, not assuming innateness in the Chomskyan sense does not consequently mean that domain-specificity must also be dismissed. Although innateness and domain-

[5] This will be discussed exhaustively in Chapters 2 and 3.

specificity are frequently viewed as inherently linked, there is no reason to assume that this is indeed the case (Khalidi, 2001)

Section 1.2.2 showed that domain-specificity and encapsulation are frequently invoked as the only way to avoid inefficient processing. However, the view that domain-specificity is not necessarily related to encapsulation or to modularity is supported by Atkinson and Wheeler (2004), who carefully dissect arguments in support of domain-specificity. Crucially, they tackle the frame problem, which is used not only by Fodor (1983) but also by Tooby and Cosmides (1992) to rule out domain-general mechanisms as a genuine possibility. The frame problem refers to cognitive mechanisms (or modules) being encapsulated to ensure efficient processing. However, as Atkinson and Wheeler point out, assuming informationally encapsulated cognitive resources does not leave domain-specificity as the only option. In their view, it is entirely plausible that any domain-general mechanism shaped by evolution will and can only be informationally encapsulated to ensure efficiency. Thus, domain-specificity is not a logical consequence of encapsulation or the frame problem. The identification of domain-specific cognitive mechanisms therefore does not (as opposed to what Fodor would argue for) shed any light on the notion of modularity.

For the purposes of the work presented in this thesis, domain-specificity will be considered in line Atkinson and Wheeler (2004). They outline that there are, in theory, four plausible options for explaining how the mind might processes stimuli (Table 1.4).

Table 1.4: The nature of the information (domain-general or domain-specific) considered in combination with the cognitive mechanisms (domain-general or domain-specific) leads to four potential information-mechanism pairings. Taken from Wheeler and Atkinson (2004). In the present context, the term 'mechanism' refers to the learning device, and the term 'information' regards the information the learning device draws from during analysis.

		Information	
		Domain-general	**Domain-specific**
Mechanisms	**Domain-general**	1. Domain-general mechanisms coupled with domain-general information	2. Domain-general mechanisms coupled with domain-specific information
	Domain-specific	3. Domain-specific mechanisms coupled with domain-general information	4. Domain-specific mechanisms coupled with domain-specific information

Firstly, general cognitive resources could be responsible for processing materials across a number of domains. Secondly, general cognitive resources could be recruited for the computation of domain-specific stimuli. Furthermore, domain-specific mechanisms could be employed for both domain-general (option 3) as well as domain-specific (option 4) information. Whereas evolutionary psychologists, such as Tooby and Cosmides (1992), assume domain-specific mechanisms are associated with domain-specific information, Atkinson and Wheeler see no in-principle reason why options 2 and 3 are largely ignored in the literature.

Importantly, Atkinson and Wheeler's more refined view of domain-specificity emphasises a crucial point: Assuming a specific learning mechanism has been identified, which has been shown to operate both in the linguistic and in the non-linguistic domain, then there are two fundamentally different possible explanations for this finding. 1. There could be one underlying mechanism, which functions across domains (Atkinson and Wheeler's option 1). 2. It could be due to separate computational systems, which happen to be identical, yet are positioned in different domains. This is in line with Atkinson and Wheeler's option 3, where domain-specific mechanisms draw from domain-general information. So, if one system was replicated in another domain, then it would also come across as a domain-general ability. To make this point even clearer: The ability to track unit-to-unit transitional probabilities is not domain-specific as it has been shown to operate in the linguistic (Saffran, Aslin & Newport, 1996) as well as in the non-linguistic domain (Kirkham, Slemmer & Johnson, 2002). In line with Atkinson and Wheeler, this could show that either a domain-general ability draws from domain-general information (option 1), or that the statistical learning tool has been replicated, and there are independently-functioning systems, which draw from domain-general information (option 3). There are thus different levels of explanation when it comes to domain-specificity.

Since Atkinson and Wheeler's approach to domain-specificity allow for a more refined analysis, the findings on domain-specificity/ -generality in this thesis will be interpreted in line with the four possibilities presented above (Table 1.4) rather than in terms of encapsulation and modularity. This will be of specific importance in Chapters 5 – 7.

1.5 Précis of the thesis

The goal of the present thesis, goes beyond stream segmentation based on non-adjacent regularities. Specifically, the experiments presented and discussed in Chapters 2 – 6 aim to explore the mechanisms employed in the detection of non-adjacent dependencies on the syntax level (i.e. between words) and to investigate whether and to what extent these mechanisms operate in the non-linguistic domain. To this end, the grammar employed by Gómez (2002), which is shown in Table 1.3 (p.35), has been used as the basis for all AGL experiments throughout the empirical work presented here.

Chapter 2 presents an experiment whose aim was to replicate Gómez's (2002) findings in the visual modality. Across two conditions the impact of the Gestalt Principle of Similarity is investigated, and the findings support the idea that featural similarity between relevant units is not a requirement for the detection of non-adjacencies in the linguistic domain. Chapter 3 explores whether this finding is robust enough to enable the identification of non-adjacencies across linguistic elements even when they are obscured. The results go against the general consensus, which claims that in an artificial language learning paradigm, non-adjacencies need to be highlighted in order to be learnt (Newport & Aslin, 2004).

Chapter 4 focuses on domain-specificity by contrasting learnability of non-adjacent dependencies across three non-linguistic conditions. In visually presented sequences, the internal structure of the non-linguistic materials was manipulated, allowing us to draw conclusions regarding stimulus simplicity. The results show that non-adjacent dependencies in non-linguistic sequences are detected when the relevant units of analysis are highlighted, i.e. the relevant units of analysis are instantiated as simple shapes. Crucially, when the non-linguistic materials are componential, i.e. designed to replicate the internal structure inherent to language, then non-adjacent dependencies are not learnt. The outcomes thus demonstrate that non-adjacencies are learnable across domains, but that language-specific expectations that guide people's attention towards relevant regularities do not carry over into the non-linguistic domain.

Since language is first and foremost an auditory signal, it has been suggested that sequence learning is facilitated in the auditory modality (Conway & Christiansen, 2005). Chapter 5 thus investigates modality effects in the acquisition of non-adjacencies by directly comparing performances in aurally presented linguistic and non-linguistic stimuli. The findings suggest that non-adjacencies in the linguistic domain are readily acquired in the auditory modality, yet their detection in the non-linguistic domain poses a challenge.

Chapter 6 offers further support for the notion of language-specific expectations identified in Chapter 4 by showing that pre-experimental familiarisation with componential non-linguistic patterns does not result in non-adjacencies being learnt in a subsequent grammar learning experiment.

Finally, Chapter 7 will summarise the main findings across all experiments and the consequences of the findings for the acquisition of non-adjacencies. Interesting future directions will also be explored.

Chapter 2

The acquisition of non-adjacencies in the linguistic domain

At the very core of successful language learning lies the ability to make sense of the rich structure and regularities found in language. Acquiring language is much more than detecting word boundaries and being able to access the mental lexicon, it also involves the correct tracking of units over which regularities operate within the given input. Due to the hierarchical structure of language, these regularities frequently occur at a distance. The previous chapter discussed empirical evidence showing that the acquisition of these non-adjacent dependencies is especially demanding. For this reason, the aim of this chapter is to investigate one of the cues that has been suggested to govern the detection of non-adjacent dependencies: The Gestalt Principle of Similarity. The Gestalt Principle of Similarity is an organising principle (Wertheimer, 1938), which stipulates that people automatically tend to group together units that are perceptually similar. How this relates to language, and specifically how this is relevant for the acquisition of non-adjacent dependencies will be discussed in detail below.

2.1. Experiment 1

As explicated in Chapter 1, non-adjacent dependencies describe regularities between elements separated by a number of intervening items. These dependencies are ubiquitous in natural languages, as illustrated by the subject-verb agreement below.

(1a) The socks on the floor are red.

*(1b)*The socks on the floor is red.*

The underlined words, the subject *The socks* and the verb *to be*, need to agree in number, as in (1a). Since the subject of the sentence is plural, the verb too must take the plural form or it will result in an ungrammatical sentence (1b). Non-adjacent dependencies are a defining property of natural languages as they can only be adequately captured by a PSG. Phrase structure rules are based on a hierarchical combination of constituents: In natural languages, sentences are organised in a hierarchical fashion, in which words are combined to form successively larger units to form phrases. PSGs therefore allow for non-adjacent constituents to form dependencies across intervening constituents.

Intriguingly, although non-adjacencies are a common feature of human languages, their acquisition has been shown to be governed by domain-general constraints, and one of these constraints seems to be the Gestalt Principle of Similarity (Gómez, 2002; Newport & Aslin, 2004). Newport & Aslin (2004) used an

AGL paradigm to explore the learnability of trisyllabic artificial words, whose structure was determined by either non-adjacent segments or non-adjacent syllables. Non-adjacencies between syllables were explored by using a Syllable Language, and non-adjacencies between segments were explored in two different Segment Languages. In their design, the Syllable Language followed an AXB pattern, with each category (A, X and B) containing CV syllables, where each CV syllable A reliably predicted a specific CV syllable B, with one intervening CV syllable X. For example, the A_B syllable pair *ba_te* could be completed by inserting the CV *di* from category X in the middle position, thus creating the artificial word *badite*. The additional intervening syllables *ku*, *to* and *pa* from the X category thus rendered *bakute*, *batote* and *bapate*, with the first and the third syllable representing the non-adjacency. By contrast, in the Segment Languages the trisyllabic words were formed by interleaving a frame of three segments with three different segments. In one of the Segment Languages, a frame of three consonants was interleaved with three vowels. Thus, the non-adjacent consonant frame p_g_t_ could take one of two vowel fillers [_a] [_i] [_ae] or [_o] [_u] [_e] creating the two artificial words *pagitae* and *pogute*, with the underlined consonant segments representing the non-adjacent regularities. In the other Segment Language, the vowel frame, for example _a_u_e, took the consonant fillers [p_] [g_] [t_] and [d_] [k_] [b_], rendering *pagute* and *dakube*.

In the initial training phase, participants were exposed to the artificial language, which was presented aurally as a continuous string with no cues to indicate word boundaries. For the testing phase, participants were required to discriminate words from partwords in a 2-alternative-forced-choice test. Newport and Aslin were therefore testing their participants' ability to segment the speech stream based on the transitional probabilities between non-adjacent units rather than their ability to learn non-adjacent dependencies. Partwords in their design were generated by either combining the final syllable of one word with the first two syllables of another (B_iA_jX), or by merging the final two syllables of one word with the first syllable of another (XB_iA_j). Each partword had therefore been presented during training. Successfully discriminating between words and partwords thus reflects the ability to track transitional probabilities between non-adjacent units. Newport and Aslin found that their participants were able to make the correct distinction between words and partwords in the Segment Languages only. One potential explanation they put forward is the Gestalt Principle of Similarity.

The Gestalt theory posits organising principles according to which people group together elements, the Principle of Similarity being one of them. In line with this law, stimuli are naturally grouped together according to featural proximity despite what their spatial or temporal relationship might be (Wertheimer, 1938/1944). A classic example used to illustrate the power of this similarity principle is shown in Figure 2.1.

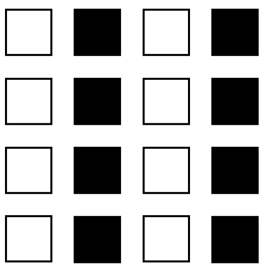


Figure 2.1: The Gestalt principle of lightness similarity. The visual is organised into separate columns of dark and light squares by the perceiver. Taken from Quinn, Burke and Rush, 1993.

Here, when perceiving the pattern in 2.1, people are guided by the Gestalt Principle of Similarity, specifically of Lightness Similarity, and they thus tend to group the squares together according to featural similarity (in this case lightness of the squares). For this reason the pattern is perceived as four columns: Two columns of white squares interleaved by columns of black squares (Quinn, Burke & Rush, 1993).

Applying this organising principle to their materials, Newport and Aslin (2004) argue that learners are more inclined to form associations between non-adjacent segments since all vowels share common features, as do all consonants. Crucially, all vowels are sonorous sounds, which are produced with a more open vocal tract, whereas consonants are segments that are produced with a more restricted vocal tract (O'Grady, Dobrovolsky & Katamba, 1996). Thus, the reason why people fail to learn the Syllable Language, according to Newport and Aslin, might be due to the non-existence of obvious similarities between the relevant units of analysis.

A closer look at the stimulus materials used in Peña, Bonatti, Nespor and Mehler's (2002) experiment reveal they too harnessed the Gestalt Principle of Similarity, albeit in a more subtle manner: Their trisyllabic AXB words were instantiated by lexical items such as [puliki], [beRaga] and [tafodu]. As Newport and Aslin (2004) point out, all the CV syllables that occur in the non-adjacent A_B dependency used by Peña et al. start with a plosive, whereas all X syllables have either an initial liquid (/l/) or fricative (/f/, /R/). Syllables from categories A and B thus share more features than they do with any of the X syllables. Participants in the

Peña et al. experiment therefore might have exploited Gestalt cues in addition to the aforementioned pause cues (see Chapter 3) to learn the non-adjacencies.

Onnis, Monaghan, Richmond and Chater (2005) replicated and extended Peña et al.'s findings. Across six experiments, Onnis et al. additionally manipulated the statistical and phonological information available to learners. In all experiments, participants were exposed to a continuous stream of synthesised speech to investigate whether the stream could be segmented using reliable transitional probabilities between non-adjacent elements. Much like in Peña et al., participants here were trained on A_1XB_1 words, with the non-adjacency holding between categories A and B. Each category included 3 CV syllables. On test, participants were required to discriminate words of the form A_iXB_i from partwords, of the form B_iA_jX or XB_iA_j. Their results indicate that phonological processing is the key to the successful detection of non-adjacencies during word segmentation. Participants showed a bias towards plosives as word onsets and they mis-segmented the speech stream when this was not the case. This bias could, however, be overcome by making use of the Gestalt Principle of Similarity: Segmentation based on non-adjacent dependencies was thus successful as long as the dependent syllables were phonologically similar. So in this case, even if they started with continuants (and not plosives), the non-adjacencies were learnable.

In a somewhat different design, Gómez (2002) investigated the acquisition of non-adjacent dependencies between artificial words (as opposed to segments and syllables). However, she also facilitated the identification of non-adjacent dependencies by harnessing the Gestalt Principle of Similarity. In her AL, the words forming the dependencies were all monosyllabic, whereas the words from category

X were bisyllabic (see Table 1.3, p.35). In doing this, Gómez included an additional cue to highlight the relevant units of analysis, which participants could have used to identify the regularities in question.

AGL work has also been carried out on the learnability of centre-embedded structures of the form $A_3A_2A_1B_1B_2B_3$ (Bahlmann & Friederici, 2006; Fitch & Hauser, 2004; Friederici, Bahlmann, Heim, Schubotz & Anwander, 2006). These A^nB^n structures produce nested dependencies and consequently non-adjacent dependencies across an arbitrary number of elements. Importantly, each unit from category A predicts a specific unit from category B (as indicated by the subscripts). These experiments are particularly relevant in the present context due to the stimulus materials used.

Specifically, as noted in the previous chapter, the CV syllables used by Fitch and Hauser (2004) for category A were perceptually distinct from the syllables used for category B as the A syllables were spoken by a female voice and syllables in category B were spoken by a male voice. Thus, all stimuli within the same category were perceptually similar. This therefore means that Fitch and Hauser included an additional cue into their experimental design: The crucial A and B elements were not merely defined in terms of the grammar, but also the similarity between syllables (female voice versus male voice), which might have assisted participants in the completion of the task. Specifically, rather than extracting the grammar underlying sequences, participants could have simply counted the syllables spoken by a female voice and by the male voice. For the FSG, then, all that was required was for each CV_{FEMALE} to be followed by a CV_{MALE}. For the PSG, it would have been sufficient for participants to realise that there had to be as many syllables spoken by a female as spoken by a male (i.e. CV_{FEMALE} - CV_{FEMALE} - CV_{FEMALE} - CV_{MALE} - CV_{MALE} - CV_{MALE}). This point was also raised by Perruchet and Rey (2005), who replicated this study and found that participants on test could not discriminate between $A_1A_2A_3B_3B_2B_1$ and $A_1A_2A_3B_3B_1B_2$ when one of the non-adjacencies was broken and the categories were not flagged by pitch differences. Thus, the way this experiment was set up means that strategies other than grammar learning could have accounted for the results (de Vries, Monaghan, Knecht & Zwitserlood, 2008).

The same is true for the materials used by Friederici et al. (2006), whose A-syllables all involved front vowels and were instantiated as {de, gi, le, ri, se, ne, ti, mi}, whereas their B-syllables ended in back vowels and comprised {bo, fo, ku, mo, pu, wo, tu, gu}. In both cases, human participants learnt the centre-embedded structure successfully. Again, Gestalt Principle of Similarity may well have assisted participants in acquiring the structures.

Marcus, Vijayan, Bandi Rao and Vishton's (1999) research on abstract rule learning of ABB, AAB and ABA patterns using synthesised speech involved non-adjacent repetitions of the same element (ABA). Marcus et al. found that 7-month old infants are capable of abstracting a rule and applying it to novel sequences, which is a powerful tool for language learning. However, non-adjacent repetitions of an element can be seen as an extreme application of the Gestalt Principle of Similarity as the repeated elements are not only similar to each other but in fact identical to each other. So again, featural similarity (or in this case sameness) may be playing an important role in participants' ability to compute non-adjacencies. An issue with regards to abstract rule learning concerns the repetition of identical units. Gómez, Gerken and Schvaneveldt, (2000) showed that adult participants are only capable of abstracting a FSG to new vocabulary when the grammar includes repeated items. Thus, so they argue, people struggle to abstract sequential relationships between elements, but excel at generalising repetition patterns. Yet in terms of language learning it is argued here that the latter ability is of limited use. Language involves a plethora of category repetitions, however, natural language very rarely involves meaningful repetition patterns of identical units.

The Gestalt Principle of Similarity has thus been frequently employed in previous research investigating non-adjacencies of different types. As described in Section 1.1, there are a myriad of examples for non-adjacent dependencies in natural languages. As Newport and Aslin (2004) point out, non-adjacent dependencies between segments do occur in natural languages. For instance, regularities between segments within words are central to the morphology of semitic languages, such as Hebrew and Arabic. Similarly, vowel harmony (to be described in detail in the following chapter), constitutes a non-adjacent dependency between vowels within words in languages, such as in Finnish and Turkish. Thus, in these instances the units participating in the non-adjacent regularities are indeed linked by featural similarity. However, there are many more types of non-adjacent dependencies found in natural languages, as discussed in detail in Chapter 1. Importantly, non-adjacencies specifically on the syntax-level (i.e. regularities between words) are not usually highlighted by additional cues (such as Gestalt cues) in natural languages. For example, in (1) *The socks on the floor are red*, the dependency holds between the noun and the verb. The regularity here therefore operates between words from two different syntactic categories, and it has been shown that words from different syntactic categories systematically differ in their phonological properties (Farmer, Christiansen & Monaghan, 2006).

To illustrate the issue even further, consider the wh-dependency presented in Chapter 1:

(6) What did John buy _?

In this case, the non-adjacency operates between the wh-word in the sentence-initial position, and its syntactically associated extraction site (see underlined elements). Here, one part of the regularity receives a null-spellout (i.e. is invisible/ inaudible). Although in accordance with Minimalism, it might be argued that the trace the moved wh-word leaves at its extraction site is in fact identical to the wh-word itself and therefore similar, it nevertheless goes against Newport & Aslin's (2004) argument. Newport and Aslin use the Gestalt Principle of Similarity to explain their results by arguing that phonological similarity facilitates the identification of non-adjacencies. However, assuming there is indeed a trace left behind at the extraction site of the moved wh-word in (6), this trace is silent, i.e. not given a spellout. Due to the null-spellout in (6), there is no phonological, audible similarity between the relevant constituents, which is problematic for theories which posit a crucial role for the Gestalt Principle of Similarity in the acquisition and processing of non-adjacent dependencies.

Non-adjacent dependencies in natural languages are thus frequently acquired without assistance from the Gestalt Principle of Similarity. For this reason, Experiment 1 was aimed at answering the following two research questions: (1) Can non-adjacent dependencies be learnt without additional Gestalt cues to make the relevant units more salient? (2) Does the inclusion of Gestalt cues improve non-adjacency learning? Since the identification of structural regularities goes beyond stream segmentation as investigated by Newport & Aslin (2004), the present experiment employed an AGL paradigm closely modelled on that employed by Gómez (2002), with a training phase and a subsequent grammaticality judgment test.

It is important to note that Gómez's experiment was conducted in the auditory modality, whereas Experiment 1 was carried out in the visual modality. There are obvious differences between language instantiated in the visual and language instantiated in the auditory modality. One crucial difference concerns the fact that speech is perceived, analysed and produced at a much earlier age than written language. Although the use of and ability for language itself is inherently human, language in its written form is not: the ability to read and write must be specifically taught and requires more deliberate effort than language acquisition itself (O'Grady, Dobrovolsky & Katamba, 1987). From an historical perspective, writing, as a symbolic representation of language, is also a much more recent development than language itself (O'Grady, Dobrovolsky & Katamba, 1987).

However, whether the origins of language are to be placed in the visual or auditory modality is to date a point of debate. The origins of language may therefore be either gestural, and thus visual, or vocal, and thereby auditory. The gestural origins of language are supported by Tomasello (2008) and Corballis (2003). Corballis argues that during hominin evolution, bipedalism allowed our predecessors to use their hands for communicative purposes. In this scenario, vocalisations were a secondary development, which emerged at a later stage as a side-effect of facial gestures. These were, according to Corballis, used as an addition to manual gestures, and proved useful when the hands became more and more occupied with tools. A crucial piece of support for the gestural origins of language are in fact primate vocalisations as these seem to be predominantly involuntary used to signal warning or mating (Corballis, 2003; Goodall, 1986). As such, from this perspective, the vocal origins of language seem inadequate to explain the emergence of language, an intentional communication tool. Dunbar (2003) argues against the gestural hypothesis as the visual modality is more restricted than the auditory modality in as much as that gestural communication in the dark is impossible and reaches fewer individuals than vocal communication. According to Dunbar (2003) and Aiello and Dunbar (1993), the evolution of fully-fledged human language was a gradual process, which may have originated in contact calls that served to maintain social contacts between individuals in a similar way that primate grooming does. This therefore allowed individuals to communicate (and, in essence, groom) a number of individuals at the same time.

Spoken and written language also differ significantly with regards to processing. The most crucial difference in the present context is the fact that speech is, by its very nature, a sequentially structured signal, whereas written language is not. Eye-tracking experiments have shown that during reading, people do not run their eyes smoothly over the text (Reichle, Rayner & Pollatsek, 2003). Rather, a combination of fixations and saccadic movements mean that readers tend to skip a number of characters or even words and pay more attention to specific lexical items. In reading English, people move their eyes from left to right most of the time, and fixate content words (Liversedge & Findlay, 2000). Importantly, 15% of the time, reading involves back-tracking and thereby moving the eyes from right to left when reading English, which can assist sentence comprehension (Liversedge & Findlay, 2000). This is not possible in the auditory modality, where the stimuli must be processed instantaneously. For the present experiment, the materials were initially presented for the following two reasons: (1) the visual modality allows for direct translation of the linguistic stimulus materials into the non-linguistic domain (this is particlualrly important for Chapter 4, and this will be discussed in detail in Chapter 7), and (2) simultaneous visual presentation allows participants to process each string thoroughly by re-reading if necessary.

Due to the obvious differences between written and spoken language discussed above, the issue of modality will be further investigated in Chapter 5.

2.1.1 Method

In order to determine the relevance of the Gestalt Principle of Similarity in the acquisition of non-adjacent dependencies, this AGL experiment included two conditions that differed with regards to available cues: The +Gestalt condition included a length cue in line with Gómez (2002), whereas the –Gestalt condition did not. The prediction was that including a Gestalt cue would facilitate the detection of the underlying grammar by highlighting the relevant units of analysis. However, it was also predicted that non-adjacent dependencies would be learnable in the absence of Gestalt cues as this is the case for many types of non-adjacent dependencies in natural languages. The grammar along with all lexical instantiations used can be found in Section 2.1.1.2.

2.1.1.1 Participants

A total of 56 adult participants were recruited from Northumbria University campus and from our research centre's pool of regular experimental participants. 24 of the participants were randomly assigned to the +Gestalt condition, and 32 to the –Gestalt condition. All participants were native speakers of English, and they either received course credit or £4.50 for participating.

2.1.1.2 Materials

The ARC Nonword database (Rastle, Harrington & Coltheart, 2002) and the MRC Psycholinguistic database[6] were utilised to generate a large number of CVC nonwords. These nonwords (under the exclusion of slang words and colloquialisms) were then individually checked for the lowest possible frequencies via Google.co.uk. The artificial words that rendered the smallest number of hits (between approximately 6,000 and 300,000) were then filtered even further. Although it was impossible to completely avoid using phonological neighbours of English words, artificial words that only differed from English words by one phonetic feature were excluded, yet artificial words that form minimal pairs with English words were not. For example, *lum*, which forms a minimal pair with *sum*, was included, whereas *zum*, which only differs from *sum* in that the initial fricative is voiced, was not included in the final set of artificial words for the –Gestalt condition. Particular care was taken when deciding on the artificial words that were to form the dependencies. None of the relevant words form minimal pairs, and – unlike in Peña Bonatti, Nespor and Mehler's (2002) design – none of them start with the same consonant. For the +Gestalt condition, an additional VC suffix was attached to the present CVC words, again ensuring low frequencies via Google and avoiding very close phonological neighbours where possible. All artificial words for both conditions can be found in Section Tables 2.1 and 2.2.

[6] http://www.psy.uwa.edu.au/MRCDataBase/uwa_mrc.htm

The grammar for both conditions was based on Gómez's (2002) grammar for the largest set size of 24 (see Table 1.3, p.35), and is displayed in Table 2.1 below.

Table 2.1: Underlying grammar for Experiment 1 (top). Lexical items used in L1 and L2 for both conditions are given below.

S_1 → aXd

S_2 → bXe

S_3 → cXf

-Gestalt condition		+Gestalt condition	
L1	L2	L1	L2
a → lum d → fip b → zel e → pof c → vok f → gam	a → nis d → huk b → jad e → zin c → fet f → gos	a → lum d → fip b → zel e → pof c → vok f → gam	a → nis d → huk b → jad e → zin c → fet f → gos
X → {*fet, fub, fum, gos, huk, hup, jad, jeg, lek, lep, lig, lof, lud, nis, nug, nup, pif, pir, taf, vam, vek, zec, zin, zog*}	X → {*fip, fub, fum, gam, hup, jeg, lek, lep, lig, lof, lud, lum, nug, nup, pif, pir, pof, taf, vam, vek, vok, zec, zel, zog*}	X → {*fetac, fubal, fumox, goseg, hukig, hupet, jadif, jegin, lekiv, lepod, ligop, lofuz, ludem, nisur, nugom, nupaf, pifar, piruk, tafep, vamex, vekas, zecid, zinev, zogik*}	X → {*fipul, fubal, fumox, gamuc, hupet, jegin, lekiv, lepod, ligop, lofuz, ludem, lumot, nugom, nupaf, pifar, piruk, pofus, tafep, vamex, vekas, vokaz, zecid, zelon, zogik*}

Both conditions included two Languages, L1 and L2, which differed in their assignment of words to categories in order to control for arbitrary preferences for specific lexical items. The artificial words forming the non-adjacent dependencies were all monosyllabic CVC words, and the non-adjacencies were identical across both conditions.

The crucial manipulation in this experiment involved category X. In the –Gestalt condition, all 24 X items were instantiated as CVC words, thereby exhibiting no obvious featural differences that would distinguish them from the words making up the dependencies. The +Gestalt condition replicated Gómez's experiment by using bisyllabic X elements. An example string for the –Gestalt condition is *lum fet fip*, with all items being monosyllabic CVC words. The equivalent string in the +Gestalt condition is *lum fetac fip*. The bisyllabic words from the X category in this condition therefore represented instantiations of the Gestalt cue, which might highlight the non-adjacencies between monosyllabic elements.

Stimuli in both conditions were presented visually, on a white computer screen. Since this experiment investigated non-adjacencies on the syntax-level, each CVC word within one sequence was separated from the next by a space. Each sequence was presented for 2500ms, with 1000ms pauses (blank screen) between each string (see Figure 2.2).

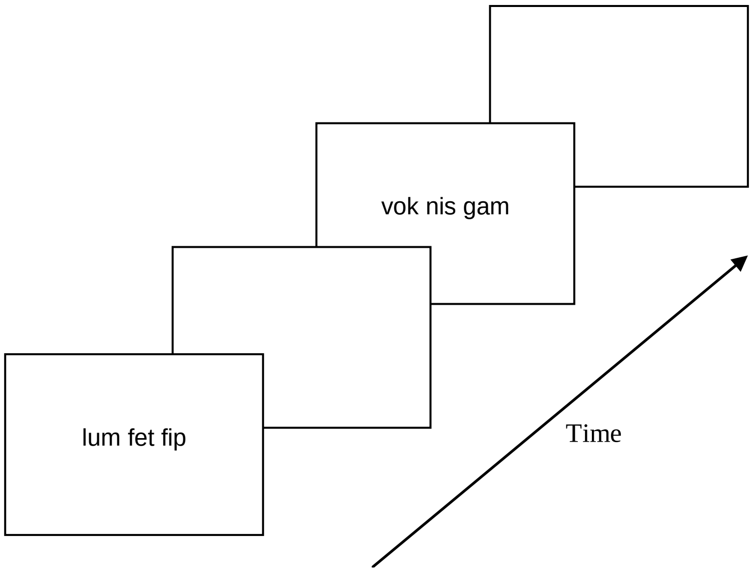


Figure 2.2: Schematic illustration of the simultaneously presented strings (black print on a white background), with each string being separated by a 1000ms presentation of a blank screen.

Since it has been shown that the acquisition of non-adjacencies in the linguistic domain can be challenging and subject to a number of different constraints, such as variability (Gómez, 2002; Onnis, Christiansen, Chater & Gómez, 2003; Onnis, Monaghan, Christiansen & Chater, 2004), pause cues (Peña et al., 2002) and processing space (Santelmann & Jusczyk, 1998; Höhle, Schmitz, Santelmann & Weissenborn, 2006), all three artificial words of each sequence for this experiment were presented simultaneously.

Saffran (2002) proposed that simultaneous presentation of materials facilitates the detection of structural regularities, which therefore suggests that visual, simultaneous presentation should increase participants' chances of learning the non-adjacencies.

The experiment was designed using the software package Slide Generator[7].

2.1.1.3 Procedure

In the initial training phase, which lasted approximately 20 minutes, participants were either exposed to L1 or L2 (with an equal number of participants taking part in both). In both languages, they viewed a total of 216 sequences, since each of the 24 X elements appeared in each of the three dependencies three times (24 x 3 x 3). Sequences were presented in random order. Participants were merely asked to pay careful attention while a large number of sequences, consisting of three made-up words, appeared on the computer screen. They were informed that the training would be followed by a test involving these sequences. There were two opportunities for participants to take a brief break during the training phase.

[7] http://www.psy.plymouth.ac.uk/research/mtucker/slidegenerator.htm

Before participants continued with the testing phase, they were told that the sequences from the training followed a specific rule, and that for each sequence that appeared on the screen during the test, they would have to decide whether or not it followed the same rule as the strings from the training. Responses were indicated by key press with their dominant index finger. The V and B keys on the keyboard served as "yes" and "no" keys and were therefore marked with either "Y" or "N". Assignment of "Y" and "N" to the keys was counterbalanced across participants. Participants were advised to rest their finger on the table, at an equal distance to both the "Y" and "N" keys. Although participants had a total of 10s to make their decision for each string, they were advised to respond as quickly as possible.

Gómez (2002) did not test to see whether her participants were able to apply the extracted grammar to sequences which were grammatical yet contained novel words: All test sequences in her experiment were either grammatical and familiar (i.e. presented during training), or ungrammatical and therefore novel. However, since generalising to sequences involving novel words is a defining feature of language and thus a crucial aspect of language learning, participants in the present experiment were tested on sequences containing both familiar and novel X items, much like in Onnis, Monaghan, Christiansen and Chater's (2004) design. Thus, in the present experiment, half of the X items participants were familiarised with during training were replaced with entirely unfamiliar X items, thereby eliminating the possibility of participants merely memorising individual sequences. In both conditions, L1 and L2 were divided into sub-groups, a and b. The sub-groups differed with regard to which X elements were replaced by novel ones, and with regard to the illegal endpoint for the grammatical violations (see Table 2.2, below, for both sub-groups in the +Gestalt and –Gestalt conditions). Grammatical test strings followed the grammar shown in Table 2.1 (p.59).

Table 2.2: Grammatical violations used in the testing phase for versions a and b in both L1 and L2 (top). Familiar (i.e. encountered in the training phase) and novel (i.e. never before seen) words for the +Gestalt and –Gestalt conditions. The same lexical items were used for both grammatical and ungrammatical sequences.

L1		L2	
Version a	Version b	Version a	Version b
*aXf	*aXe	*aXf	*aXe
*bXd	*bXf	*bXd	*bXf
*cXe	*cXd	*cXe	*cXd
X_{+Gest_fam} → *zinev, ligop, goseg, vamex, vekas, fetac, lekiv, nupaf, piruk, fubal, ludem, hukig*	X_{+Gest_fam} → *nisur, lofuz, zokig, jadif, tafep, jegin, zecid, lepod, fumox, nugom, hupet, pifar*	X_{+Gest_fam} → *ligop, lofuz, tafep, vekas, lekiv, nupaf, piruk, fubal, ludem, fipul, zelon, vokaz*	X_{+Gest_fam} → *pifar, zogik, vamex, jegin, zecid, lepod, fumox, hupet, lumot, pofus, gamuc, nugom*
X_{+Gest_nov} → *habec, natuf, jafen, zepak, jevat, pesir, hinug, jukel, rudil, zomun, noseg, vugap*	X_{+Gest_nov} → *kamut, narel, larof, wemic, tegor, kebam, jufis, musov, dovad, vogud, gubip, gedok*	X_{+Gest_nov} → *habec, natuf, jafen, zepak, jevat, pesir, hinug, jukel, rudil, zomun, noseg, vugap*	X_{+Gest_nov} → *kamut, narel, larof, wemic, tegor, kebam, jufis, musov, dovad, vogud, gubip, gedok*
X_{-Gest_fam} → *zin, lig, gos, vam, vek, fet, lek, nup, pir, fub, lud, huk*	X_{-Gest_fam} → *nis, lof, zog, jad, taf, jeg, zec, lep, fum, nug, hup, pif*	X_{-Gest_fam} → *lig, lof, taf, vek, lek, nup, pir, fub, lud, fip, zel, vok*	X_{-Gest_fam} → *pif, zog, vam, jeg, zec, lep, fum, hup, lum, pof, gam, nug*
X_{-Gest_nov} → *hab, nat, jaf, zep, jev, pes, hin, juk, rud, zom, nos, vug*	X_{-Gest_nov} → *kam, nar, lar, wem, teg, keb, juf, mus, dov, vog, gub, ged*	X_{-Gest_nov} → *hab, nat, jaf, zep, jev, pes, hin, juk, rud, zom, nos, vug*	X_{-Gest_nov} → *kam, nar, lar, wem, teg, keb, juf, mus, dov, vog, gub, ged*

In the test, each of the three dependencies occurred with 12 familiar and 12 novel X elements, as did each violated dependency, rendering a total of 144 test items ((3 dependencies + 3 violated dependencies) * 24 X elements).

The testing phase usually lasted approximately 25 minutes. Importantly, the testing phase here was much more comprehensive than the testing phase in Gómez's design. For her experiment involving adult participants, the testing phase required them to make grammaticality judgements for a total of 12 strings, 6 grammatical and 6 ungrammatical. The present experiment, with 144 test items would thus result in a much more reliable assessment of people's abilities with regard to the extraction of non-adjacent dependencies.

2.1.2 Results

Table 2.3 reports the mean percentage endorsements for grammatical and ungrammatical sequences, involving familiar and novel X words across the +Gestalt and –Gestalt condition.

Table 2.3: Mean percentage endorsements for grammatical and ungrammatical strings, involving familiar and unfamiliar X items, across both conditions.

Condition	Grammaticality	Familiarity	% endorsements	SD
+Gestalt	grammatical	familiar	74.65	20.07
		novel	62.38	25.56
	ungrammatical	familiar	44.68	31.70
		novel	30.44	23.32
-Gestalt	grammatical	familiar	67.80	16.39
		novel	51.65	23.75
	ungrammatical	familiar	45.31	23.13
		novel	30.47	19.02

In order to check whether Subgroups (L1a, L1b, L2a, L2b) should be included as a factor in the subsequent analysis, two One-Way ANOVAs were conducted on the overall correct responses for both conditions. First, a One-Way ANOVA contrasting performances in L1a, L1b, L2a, L2b was carried for the +Gestalt condition. An alpha level of $p < .05$ was employed for this ANOVA and for all following analyses unless stated otherwise[8]. Since no significant differences were found for correct responses, $F(3, 20) = 1.29$, $p = .305$, all +Gestalt sub-groups were pooled together for further analyses. Similarly, a One-Way ANOVA for the –Gestalt condition resulted in a non-significant effect, $F(3, 28) = .185$, $p = .906$, and consequently all these sub-groups were combined for further analyses.

A 2 x 2 x 2 ANOVA was carried out to contrast the percentage of test sequences participants endorsed across the +Gestalt and –Gestalt conditions, with Condition as a between-subjects factor, and Familiarity with X (familiar, novel) and Grammaticality (grammatical, ungrammatical) as within-subjects factors. The full ANOVA table can be found in Appendix B (Table 1). The analysis resulted in two main effects.

The main effect for Grammaticality, $F(1, 54) = 28.49$, $p < .001$ is due to the fact that participants accepted significantly more grammatical (64.12%) than ungrammatical (37.72%) sequences, regardless of all other factors. This shows that overall, participants succeeded at extracting the non-adjacent dependencies (see Figure 2.3A).

[8] For the present ANOVAs, as for all subsequent ANOVAs presented in this thesis, should the assumption of equal variances be violated, the significant Levene's test will be reported along with the Welch F ratio.

The main effect for Familiarity, $F(1, 54) = 48.54$, $p < .001$, demonstrates that, regardless of the grammaticality of the sequences, participants favoured strings containing familiar X units as they endorsed 58.11% familiar sequences as opposed to 43.74% of strings with novel X elements (see Figure 2.3B).

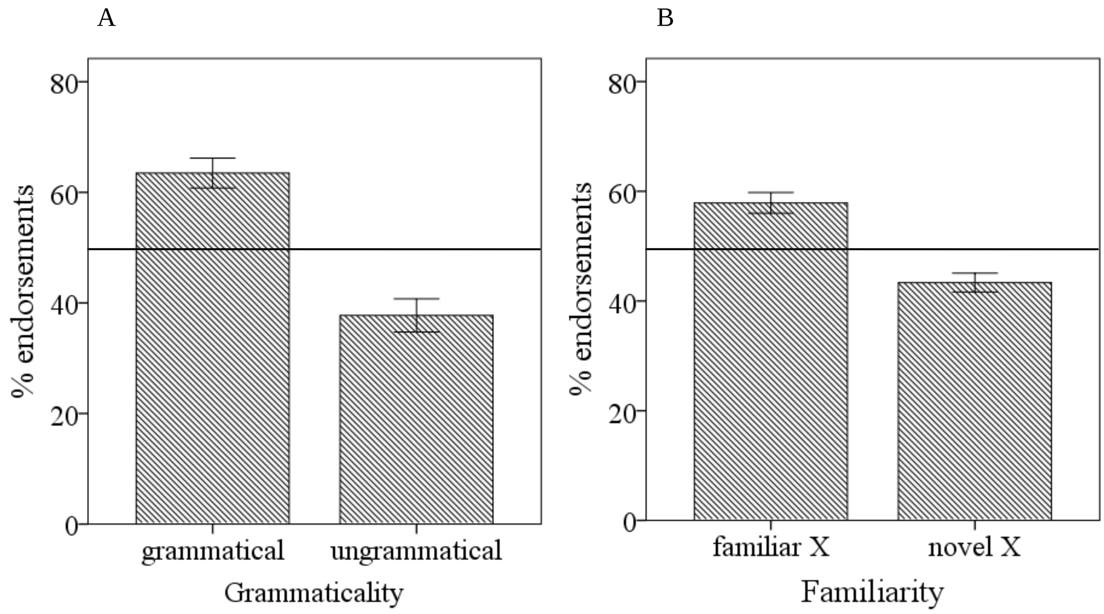


Figure 2.3: Main effect for Grammaticality (A) and Familiarity (B), with standard error bars (+/- 1 SE). The horizontal line indicates 50% chance performance. In both conditions, participants endorsed significantly more grammatical strings than ungrammatical strings (A), as well as significantly more familiar than unfamiliar strings (B).

The overall ANOVA revealed no further significant effects or interactions. There was thus no significant between-subjects effect (i.e. no effect for Condition), $F(1, 54) = 2.02$, $p = .161$, indicating that the overall endorsement rates across both conditions do not differ significantly. Importantly, there was also no significant Grammaticality x Condition interaction ($F(1, 54) = .85$, $p = .360$). This, in combination with the main effect for Grammaticality shows that participants in both conditions successfully discriminated between grammatical and ungrammatical sequences and therefore successfully extracted the non-adjacent dependencies. As shown in Figure 2.4 (p.68), participants who received assistance from the Gestalt

Principle of Similarity accepted slightly more grammatical strings compared with participants who had no additional Gestalt cue, yet this difference (Grammaticality x Condition interaction) was non-significant.

The overall analysis is aimed at determining whether the underlying grammar was learnt. In order to confirm the main effect for Grammaticality, two separate paired-samples t-tests were conducted to contrast percentage endorsements with grammatical strings against percentage endorsements with ungrammatical strings, separately for both conditions. Since two t-tests were conducted, a manually corrected alpha-level of $\alpha = .05/2 = .025$ was employed. The results from the t-tests show that endorsement rates between grammatical and ungrammatical sequences differed significantly across both conditions (see Table 2.4).

Table 2.4: Mean percentage endorsements for grammatical and ungrammatical strings, along with results from the paired samples t-tests, for sequences containing familiar and unfamiliar X items.

Condition	Grammaticality	% endorsements	Statistic
+Gestalt	grammatical	68.51% (21.72 SD)	t(23) = 3.83, p = .001
	ungrammatical	37.56% (26.23 SD)	
-Gestalt	grammatical	59.72% (18.46 SD)	t(31) = 3.61, p = .001
	ungrammatical	37.89% (19.75 SD)	

As a follow-up, it was important to ensure that endorsement rates for grammatical and ungrammatical were significantly different from chance, as this would further confirm that the underlying non-adjacencies were indeed extracted. Thus, one-sample t-tests were run on the endorsement rates with grammatical and ungrammatical sequences. Participants in the +Gestalt condition, $t(23) = 4.18$, $p < .001$ (68.52%, 21.72 SD), as well as in the –Gestalt condition, $t(31) = 2.98$, $p = .006$

(59.72%, 18.46 SD) accepted significantly more than 50% of grammatical strings. Importantly, endorsements for ungrammatical strings were significanlty below chance performance in the +Gestalt, t(23) = 2.32, p = .029 (37.56%, 26.23 SD) and in the –Gestalt, t(31) = 3.47, p = .002 (37.89%, 19.75 SD). These results therefore confirm that the underlying non-adjacencies were identified equally well in both conditions (see Figure 2.4).

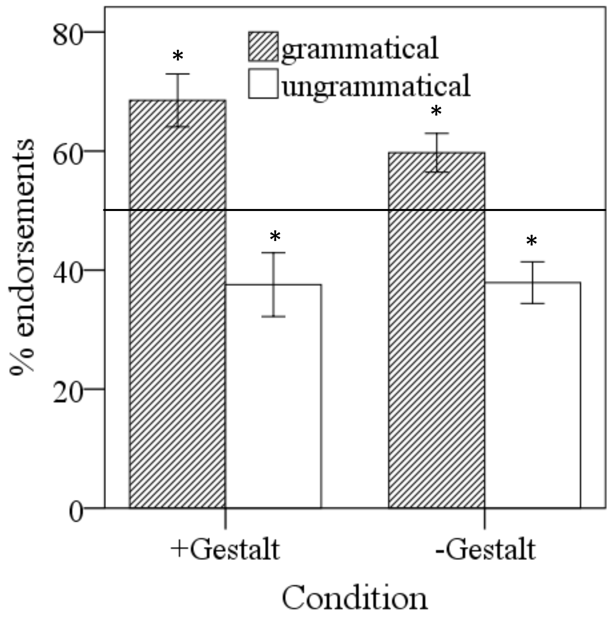


Figure 2.4: Percentage endorsements for grammatical and ungrammatical sequences across both conditions, with standard error bars (+/- 1 SE). The horizontal line across indicates 50% chance performance. Asterisks indicate performance significantly different from chance.

The main effect for Familiarity clearly demonstrates the importance of including novel X elements in the test sequences. The present results indicate that familiarity with the surface form of the test items affects participants' ability to base their grammaticality judgements exclusively on the grammaticality of the test sequences. Abstracting the non-adjacency rule and applying it to novel test strings thus poses a genuine challenge for learners.

2.1.3 Discussion

The correct identification of units over which regularities operate is a crucial aspect of language learning. Regularities that hold between words separated by a number of intervening words are a defining feature of the hierarchical structure underlying natural languages. Yet despite this, the acquisition of non-adjacent dependencies has previously been found to be a challenge for language learners, and it has therefore been suggested that a number of additional cues may facilitate the detection of non-adjacencies, the Gestalt Principle of Similarity being one of them (Newport & Aslin, 2004).

The purpose of this experiment was therefore to investigate the importance of the Gestalt Principle of Similarity in the acquisition of non-adjacent dependencies. Specifically, it aimed at investigating whether non-adjacent dependencies could be learnt without assistance from the Gestalt Principle of Similarity, and whether the Gestalt cue improves learning.

If being able to track these regularities were indeed due to Gestalt cues, participants in the +Gestalt condition would be expected to significantly outperform participants in the –Gestalt condition. As illustrated in Figure 2.3A (p. 66) as well as by the absence of a significant Condition x Grammaticality interaction, this is not the case. Crucially, the absence of this interaction also demonstrates that Gestalt cues do not increase the learnability of non-adjacent dependencies between words. Participants in both conditions reliably distinguished grammatical from ungrammatical sequences (see Table 2.4, p.67), showing that the underlying grammar was detected in both conditions. Thus, the salient length cue as used by Gómez (2002) is not an essential constraint governing regularities between non-adjacent words. The present findings therefore challenge previous assumptions that the detection of non-adjacencies is constrained by the Gestalt Principle of Similarity (Newport & Aslin, 2004). At the same time, however, these findings are perhaps unsurprising, given that, as discussed earlier, in natural languages, there are not always obvious Gestalt cues that assist learners in the detection of non-adjacent regularities between words, and thus assuming them to be a requirement seems counter-intuitive.

There are two crucial differences between the present experiment and Newport and Aslin's as well as Peña, Bonatti, Nespor and Mehler's (2002). Firstly, the present experiment tested people's capacity to extract non-adjacent dependencies whereas Newport and Aslin's and Peña et al.'s experiments were segmentation tasks. It is thus possible that the cognitive mechanism employed in learning non-adjacencies is distinct from the one recruited for stream segmentation. Secondly, and more probably, Newport and Aslin and Peña et al. only used a very limited set size for their X items, which did not offer sufficient variability between adjacent elements to disrupt element-to-element transitional probabilities and consequently guide people's attention toward more reliable structures within the input, i.e. the non-adjacencies. Thus, additional salient cues, such as Gestalt cues, were required for participants to detect the underlying structure in their experiments. Non-adjacent dependencies in natural languages occur across a large variety of intervening words, and thus it seems reasonable to replicate this fact in the experimental design. Non-adjacencies between words can be extracted when high variability of the intervening element is ensured (Gómez, 2002).

However, despite the grammar being learnt, the present findings further support the claim that non-adjacencies are difficult to detect. Participants viewed all three items for each sequence simultaneously, and there were obvious gaps between individual words. On the surface of it, this would seem to facilitate the detection of the non-adjacencies as this presentation immediately offers the correct level of analysis, i.e. regularities occurring between words. However, even with the inclusion of this gap cue, overall endorsement with grammatical strings in the +Gestalt condition was around 70%, and around 60% in the –Gestalt condition. Although these endorsement rates are significantly above chance level, they are not as high as in Gómez's experiment, who reports a mean percentage endorsements of 100 for grammatical strings for adult participants in set size 24. A possible explanation for the discrepancy found between endorsements with grammatical sequences in this experiment and Gómez's findings might be due to modality or mode of presentation. There is a possibility people might be equipped with particularly strong learning biases for aurally presented sequences, which do not carry over into the visual modality and simultaneous presentation. The issues of mode of presentation and modality-specificity are addressed in Chapters 4 (specifically Experiment 4) and 5.

Furthermore, the introduction of novel X elements to the testing phase significantly affected performance. As can be seen in Figure 2.3B (p.66), participants were generally more inclined to endorse with familiar sequences. In the +Gestalt condition, people accepted grammatical strings approximately 30% more often than ungrammatical strings, regardless of whether the X item was familiar or novel (see Table 2.3, p.64). This pattern of results resembles the –Gestalt condition, where the difference between endorsements with grammatical and ungrammatical sequences also remains constant, at approximately 20%, regardless of familiarity with X.

Crucially, however, even under the best possible conditions, i.e. strings containing familiar X words in the +Gestalt condition, endorsements for grammatical test items still only reaches 75%. The inclusion of novel X items was thus an important manipulation as it aimed at testing participants' ability to extract the non-adjacent regularities and to apply the underlying grammar to unfamiliar sequences. What this manipulation revealed was that the acquisition of non-adjacencies in the present design truly was a challenge for the learners. These findings therefore show that in AGL experiments, testing people's ability to generalise to new instantiations is relevant to assessing their ability to extract the underlying grammar.

2.2 Concluding remarks

In conclusion, this experiment supports the view that the Gestalt principle of Similarity does not act as a constraint on non-adjacency learning in the visual modality, using linguistic stimuli. Non-adjacencies in natural languages can take many different forms, and hold between elements that do not display featural similarity on the phonetic level, and thus these findings conform with the expectations laid out earlier. The present experiment thus challenges findings on non-adjacency learning to date (Gómez, 2002; Newport & Aslin, 2004; Peña, Bonatti, Nespor & Mehler, 2002) as it found that (1) non-adjacencies between words can be acquired without assistance from the Gestalt cue, (2) Gestalt cues do not improve learnability of non-adjacent dependencies. Experiment 1 also shows that tracking remote dependencies between words is a demanding task, specifically when participants are required to apply the learnt grammar to novel sequences.

However, since the visible gaps between individual words might have rendered the relevant units of analysis more salient in these experiments, thus reducing the need for a Gestalt cue, the following chapter will investigate people's ability to track on-adjacent regularities in the visual modality when there are no interword breaks.

Chapter 3

The role of break cues in non-adjacency learning in the linguistic domain

The previous chapter showed that regularities between non-adjacent monosyllabic words are reliably detected in simultaneously presented sequences in the visual modality. However, the materials in Experiment 1 displayed interword breaks. These breaks separated each CVC word from the next in each string, which could have acted as a helpful cue to participants on test. For example, in the sequence *lum fet fip*, the breaks between *lum* and *fet* and between *fet* and *fip* may have indicated that the relevant structures underlying the sequences are to be found between words (as opposed to between graphemes, for instance). Thus, the sequences in Experiment 1 were presegmented in a way that may have assisted participants in determining what level of analysis the crucial units are on. The aim of this chapter is therefore to explore the limitations of the human capacity to detect non-adjacencies in the linguistic domain by manipulating these break cues.

As mentioned in the previous chapter, similarly to Saffran, Aslin and Newport (1996), Newport and Aslin (2004) also investigated adults' performance in discriminating words from partwords, yet with the underlying structural regularity in this case being non-adjacent dependencies. Their findings (non-adjacencies between segments within an artificial word are learnable, non-adjacencies between syllables of an artificial word are not learnable) are discussed in light of Gestalt cues in Experiment 1. This issue regarding the acquisition of non-adjacencies within words will be revisited in this chapter, and examined from a different perspective.

3.1. Experiment 2

Breaking the fluent speech stream into meaningful units is an enormous challenge for infants acquiring their native language or even for L2 learners. When exposed to an entirely novel language (as is the case for infants acquiring L1 and for L2 learners), the absence of obvious pauses between words in speech obscures boundaries between individual elements, which makes the correct identification of words difficult. As mentioned in Chapter 1, Saffran, Aslin and Newport (1996) showed that infants are capable of harnessing the statistical structure underlying an unsegmented, aurally presented AL to determine word boundaries. Transitional probabilities between adjacent speech sounds thus act as a reliable cue in word segmentation. Similarly, non-adjacent regularities between CV syllables can also be harnessed for speech stream segmentation (Onnis, Monaghan, Richmond & Chater, 2005).

However, the mechanisms involved in discriminating words from partwords based on probabilistic transitions between certain elements may be fundamentally different from extracting non-adjacent regularities between specific units in an AGL paradigm. For one, speech stream segmentation is about identifying words, whereas the paradigms, such as the one employed in Experiment 1 and in the present experiment, investigate the extraction of non-adjacencies on the syntax-level. Moreover, for the stream segmentation tasks, participants are usually required to discriminate between words and partwords (or sometimes non-words). By contrast, the AGL paradigm used throughout this thesis involves extraction of dependencies between a total of 6 units, and this knowledge is then directly tested in the grammaticality judgments. Importantly, Experiment 1 (as all of the following AGL experiments presented in the following chapters) also tested participants' ability to abstract the internalised regularities to sequences containing novel X items, whereas segmentation tasks typically only investigate people's knowledge of the AL encountered during training.

Despite these obvious differences, it is important to note that in segmentation tasks involving non-adjacencies, pause cues have been shown to play a significant role (Peña, Bonatti, Nespor & Mehler, 2002). Mueller, Bahlmann and Friederici (2008) further explored the Peña et al. finding by using ERPs. In an aural AGL paradigm, they trained and subsequently tested participants on AXB sequences. In this experiment, each category (A, B and X) contained artificial CVCV words, and on test participants were required to make grammaticality judgements on sequences of either the form AXB (grammatical) or AXX[9] (ungrammatical). In doing this, Mueller et al. investigated people's ability to identify structural, non-adjacent regularities between word-like units rather than people's ability to segment a linguistic stream based on non-adjacent regularities between syllables of one word as

[9] Note that these two X elements were not identical.

was done by Peña et al. A crucial manipulation in the Mueller et al. design involved pause cues, with one condition containing a 500ms pause between sequences, thereby clearly indicating the boundaries between each AXB string, and with one condition lacking pauses and thus consisting of one continuous stream. Their findings indicate that the pause cues were required for participants to fully acquire the non-adjacencies. So, in their experiment, it was not individual words that were highlighted by pause cues. Rather, Mueller et al. argue that people require some indication as to where phrase boundaries (i.e. boundaries between each AXB sequence) lie in order to correctly identify non-adjacent dependencies.

Further evidence for the importance of pause cues comes from research on linguistic disfluencies. Silent pauses can have a number of physiological causes, but from a cognitive perspective, silences in speech production are frequently associated with problems of lexical retrieval (Butterworth, 1980; Kircher, Brammer, Levelt, Bartels & McGuire, 2004). However, silent pauses also have an effect on the listener's processing. Zellner (1994) suggests that silent pauses produced by fluent speakers will usually occur at prosodic or syntactic boundaries, and occupy a "beacon" (p.47) position within the speech stream. The purpose of these interword breaks would be to assist the speaker as well as the listener in segmenting the speech stream and thereby facilitating comprehension. Language, when presented aurally, is a temporally sequential signal, and silent pauses disrupt this temporal structure and disconnect the following word from the preceding context. Using an ERP paradigm, MacGregor, Corley and Donaldson (forthcoming) have shown the effects these disruptions have on listeners, and particularly two findings are relevant here: (1) silent pauses attenuate the N400 effect and (2) they boost recognition memory for words directly following the silences. The N400 effect refers to a brain signature, which indicates the difficulty people face when trying to integrate an unpredictable word. MacGregor et al. hypothesise that silent pauses and the disruption to the speech signal cause the hearer to prepare for an unexpected target word and thereby reduced the N400. This demonstrates that pauses in the speech stream have significant consequences for the listener.

Although pauses in the auditory modality differ perceptually from breaks in the visual modality, the interword breaks found in the materials for Experiment 1 could nevertheless have been helpful to the participants. Not only could they have given an indication of the fact that the important regularities occurred on the phrase structure level (i.e. between words as opposed to between graphemes, for example), but also, in line with MacGregor et al., there is a possibility that these breaks increased the salience of the following word.

The aim of the present experiment was therefore to explore the limits of non-adjacency learning by specifically manipulating the break cues in the materials used for the –Gestalt condition in Experiment 1. In the One-word condition, interword breaks were eliminated, thus creating one CVCCVCCVC word (as opposed to a CVC CVC CVC string as in Experiment 1). According to the literature discussed above, it would make sense to assume that eliminating the breaks would reduce people’s ability to detect individual words as easily and consequently may affect the extraction of non-adjacencies, since the relevant units over which the non-adjacencies operate are less obvious. The present experiment further included a Nine-letter condition, which involved equal breaks between each grapheme (i.e. C V C C V C C V C strings). The materials in this condition also eliminated the helpful pauses found in the –Gestalt condition of Experiment 1, without making the surface form of strings appear as single words. The research questions to be answered in Experiment 2 thus are: (1) Does the elimination of helpful interword breaks affect people's ability to extract non-adjacencies? and (2) Does the insertion of unhelpful inter-grapheme breaks make the detection of non-adjacent dependencies harder?

3.1.1 Method

3.1.1.1 Participants

48 Northumbria University students participated in this experiment. The 39 adult females and 9 adult males were randomly assigned to one of two conditions (One-word, Nine-letter), for later comparisons with the –Gestalt condition. All were native speakers of English, and they received either £4.50 or course credit for taking part.

3.1.1.2 Materials

For this experiment, the same underlying grammar was employed as in Experiment 1, and the materials were based on the stimuli used for the –Gestalt condition. The crucial difference regarded the breaks between words. The helpful breaks between each CVC word from the –Gestalt condition in Experiment 1 were removed for the One-word condition, thereby forming CVCCVCCVC strings. For the Nine-letter condition, breaks were inserted between each grapheme. The non-adjacent dependencies still operated between the first and final CVC sequence of each string. Table 3.1 illustrates how materials for the One-word and Nine-letter conditions were derived from the –Gestalt condition.

Table 3.1: Example sequences for the One-word and Nine-letter condition, and the –Gestalt strings that served as a basis. Ungrammatical sequences are indicated by an asterisk. In all three conditions the non-adjacent dependencies involve the first and third CVC passages, i.e. lum – fip, nis – huk, fet – gos and their equivalents.

Experiment 1	**Experiment 2**	
-Gestalt	**One-word**	**Nine-letter**
lum fet fip *lum fet gam	lumfetfip *lumfetgam	l u m f e t f i p *l u m f e t g a m
nis fet huk *nis fet gos	nisfethuk *nisfetgos	n i s f e t h u k *n i s f e t g o s
fet taf gos *fet taf zin	fettafgos *fettafzin	f e t t a f g o s *f e t t a f z i n

3.1.1.3 Procedure

The procedure was identical to the one used in Experiment 1 in all respects except the stimulus materials. Participants were thus trained on 216 visually presented sequences of either the One-word or Nine-letter condition, and subsequently tested. To control for arbitrary preferences individual participants might have, they were randomly assigned to complete either Language 1 version a, Language 1 version b, Language 2 version a or Language 2 version b. See Tables 2.1 (p. 59) and 2.2 (p. 63) for lexical items used in these sub-groups. Again, the testing phase involved both 50% familiar X elements, i.e. middle CVC sequences encountered during training, and 50% entirely novel X items, in order to rule the possibility of memorisation as an explanation for correct identification of non-adjacencies. As in Experiment 1, the sequences were displayed on a computer screen as black print on a white background, each sequence was presented for 2500ms, and each string was separated from the next by a blank (i.e. white) screen, which was displayed for 1000ms. This experiment was also designed using Slide Generator.

3.1.2 Results

Table 3.2 reports the mean percentage endorsements for grammatical and ungrammatical strings, collapsed across sub-groups, for the One-word and Nine-letter conditions.

Table 3.2: Mean percentage endorsements (and standard deviations) for grammatical and ungrammatical strings, containing familiar and novel CVC chunks as X items, across both conditions.

Condition	Grammaticality	Familiarity	% endorsements	SD
One-word	grammatical	familiar	75.23	17.79
		novel	60.07	30.05
	ungrammatical	familiar	42.13	29.11
		novel	25.00	21.71
Nine-letter	grammatical	familiar	62.96	12.14
		novel	50.00	18.06
	ungrammatical	familiar	58.22	15.06
		novel	41.09	18.31

Initially, a one-way ANOVA was carried out in both conditions to compare correct responses across sub-groups. In the One-word condition, the Levene's test was significant ($p = .003$), and therefore the Welch F-ratio is reported. The ANOVA revealed that overall performance across all subgroups in the One-word condition (i.e. L1a, L1b, L2a, L2b) did not differ significantly, $F(3, 8.93) = 2.47$, $p = .129$. In the Nine-Letter condition, there was also no significant difference in performance across the four subgroups, $F(3, 20) = 1.39$, $p = .276$, and for this reason the subgroups here were also collapsed for further analysis.

Percentage endorsements were submitted to a 2 (One-word vs Nine-letter condition) x 2 (familiar vs unfamiliar X) x 2 (grammatical vs ungrammatical strings) ANOVA. For the complete results from this ANOVA, see Table 2 in Appendix B. The analysis revealed two main effects and one interaction.

The two main effects here were the same as in Experiment 1. The main effect for Grammaticality, $F(1,46) = 22.12$, $p < .001$, is due to people's successful discrimination between grammatical (62.07%) and ungrammatical (41.61%) sequences (see Figure 3.1A). The main effect for Familiarity, $F(1, 46) = 44.12$, $p < .001$ (see Figure 3.1B) shows that participants overall endorsed with significantly more sequences containing familiar X chunks (59.64%) as opposed to sequences with novel X chunks (44.04%).

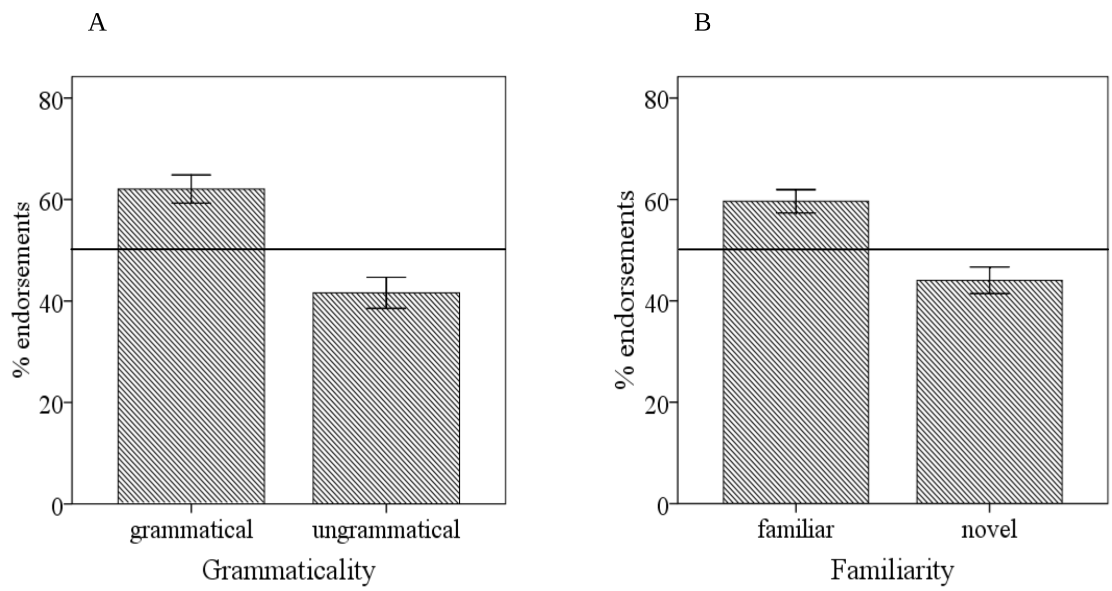


Figure 3.1: Main effect for Grammaticality (A) and Familiarity (B) and Grammaticality, with error bars (+/- 1SE). Horizontal lines indicate chance level performance (50%). Participants favoured grammatical strings over ungrammatical strings (Figure 3.1A) and strings containing familiar X sequences over strings involving novel X chunks (Figure 3.1B).

Importantly, there was also a Condition x Grammaticality interaction, F(1,46) = 9.82, p = .003 (see Figure 3.2), indicating that grammaticality judgments differed between the two conditions.

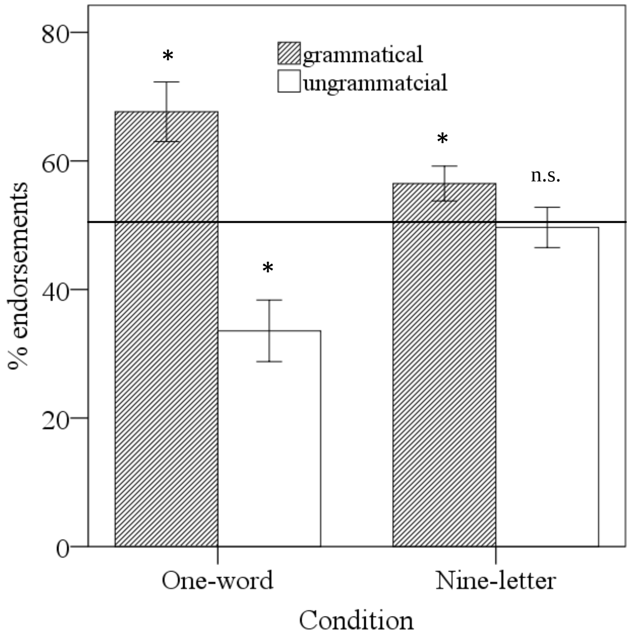


Figure 3.2: Condition x Grammaticality interaction, with error bars (+/- 1SE). Horizontal lines indicate chance level performance (50%). Asterisks indicate performance significantly different from chance.

There were no other significant main effects or interactions. This Condition x Grammaticality interaction will be further explored shortly, when contrasted against performance in the –Gestalt condition in a series of follow-up t-tests.

In order to determine to what extent break cues play a role in the detection of non-adjacent dependencies, the current findings were contrasted against the results for the –Gestalt condition from Experiment 1, which involved breaks between each CVC word. To this end, the percentage endorsements were submitted to a Stimulus-type (One-word, Nine-letter, -Gestalt) x Grammaticality (grammatical, ungrammatical) ANOVA. This analysis resulted in a main effect for Grammaticality, F(1, 77) = 33.86, p < .001, and a Stimulus-type x Grammaticality interaction, F = (2, 77) = 4.41, p = .015. Again, the main effect shows that participants overall endorsed with significantly more grammatical (61.13%, 18.88 SD) than ungrammatical (40.12%, 20.62 SD) sequences. There were no other significant effects or interactions. The complete results of this 3 x 2 ANOVA can be found in Table 3 in Appendix B.

To further explore both the Condition x Grammaticality interaction from the previous ANOVA as well as the present Stimulus-type x Grammaticality, a number of t-tests were conducted. As the key issue regards the acquisition of non-adjacent dependencies in each of the conditions (and thus for each of the stimulus types), two paired-samples t-tests comparing percentage endorsements for grammatical and ungrammatical strings across the One-word, Nine-letter and –Gestalt conditions were conducted (see Table 3.3).

Table 3.3: Mean percentage endorsements for grammatical and ungrammatical strings, along with results from paired samples t-tests for the –Gestalt, One-word and Nine-letter condition.

Stimulus-type	Example string	Grammaticality	% endorsements	Statistic
-Gestalt	lum fet fip	grammatical	59.72 (18.46 SD)	t(31) = 3.61,
		ungrammatical	37.89 (19.75 SD)	p = .001
One-word	lumfetfip	grammatical	67.65 (22.77 SD)	t(23) = 4.05,
		ungrammatical	33.57 (23.49 SD)	p = .001
Nine-letter	l u m f e t f i p	grammatical	56.48 (13.28 SD)	t(23) = 3.15,
		ungrammatical	49.65 (15.38 SD)	p = .004

Due to the fact that three t-tests were conducted, a corrected alpha-level of α = .0167 was employed. The t-tests showed that the difference between percentage endorsements for grammatical and ungrammatical strings was significant for all three stimulus types. Experiment 1 already demonstrated that the non-adjacencies were detected in the –Gestalt condition. The present results indicate that the underlying grammar was also identified in the One-word and Nine-letter conditions.

Another crucial apsect to be explored regards the difference between endorsement rates for grammatical and ungrammatical sequences, since the larger the difference between accepted grammatical and accepted ungrammatical strings, the better the discrimination between sequences based on grammaticality. Thus, comparing difference scores (i.e. percentage endorsements with grammatical sequences – percentage endorsements with ungrammatical sequences) across the conditions gives a direct indication on how well the grammar was learnt for each of the stimulus types. For this reason, the differences scores were calculated and submitted to a Tukey post-hoc analysis. This analysis demonstrated that discrimination based on the grammaticality of the sequences in the One-word condition (34.09%, 41.29 SD) was significantly better than in the Nine-letter condition (6.63%, 10.62 SD), $p = .011$, but not better than in the –Gestalt condition (21.83%, 34.26 SD), $p = .333$. None of the other pair-wise comparisons were significant.

In order to yet further investigate the Condition x Grammaticality (and simultaneously the Stimulus x Grammaticality) interaction, a number of separate one-sample t-tests were carried out in order to check whether endorsement rates differed significantly from chance. For the –Gestalt condition, this was done in Experiment 1, showing that percentage endorsements for grammatical strings were significantly above and for ungrammatical strings significantly below chance (see Figure 2.4, p. 68). The results here show that in both the One-word condition, $t(23) = 3.8$, $p = .001$ (67.65%, 22.77 SD), and Nine-letter condition, $t(23) = 2.39$, $p = .025$

(56.48%, 13.28 SD), participants endorsed with significantly more than 50% of grammatical sequences. Interestingly, however, only participants in the One-word condition successfully endorsed with less ungrammatical strings than would be expected by chance, $t(23) = 3.43$, $p = .002$ (33.56%, 23.49 SD). Participants in the Nine-letter condition, by contrast, failed to do so, $t(23) = .111$, $p = .913$ (49.65%, 15.38 SD).

The present results therefore show that participants successfully discriminated between grammatical and ungrammatical strings, and therefore extracted the non-adjacent regularities, in both the One-word and Nine-letter conditions much like in the –Gestalt condition (as indicated by the paired-samples t-tests). However, participants' competence in identifying the non-adjacencies particularly in the One-word condition was superior to the Nine-letter condition, as indicated by the Tukey post-hoc analysis and people's chance-level performance on ungrammatical strings. This difference in ability to detect non-adjacencies across the One-word and Nine-letter condition therefore explains the interactions.

3.1.3 Discussion

The results from Experiment 2 show that the detection of non-adjacent dependencies in visually presented linguistic sequences does not need assistance from obvious break cues to highlight the correct level of analysis. In Experiment 1, helpful breaks between individual CVC words could have given participants an indication that the crucial regularity for each sequence is to be found between words. Experiment 2 addressed this issue by manipulating the breaks in the sequences. The One-word condition investigated the learnability of non-adjacent dependencies within

CVCCVCCVC words, where there were no obvious interword breaks. The Nine-letter condition focused on the acquisition of non-adjacencies in C V C C V C C V C strings, with equal breaks between each grapheme. Crucially, the paired-samples t-tests indicated that the underlying grammar was acquired in both conditions (see Table 3.3, p.84). However, participants in the One-word condition significantly outperformed participants in the Nine-letter condition. This was shown by the Condition x Grammaticality interaction, and was further supported by the post-hoc analysis. Mean percentage endorsements between grammatical and ungrammatical sequences in the One-word condition differ by 34.09%, whereas endorsement rates for grammatical sequences in the Nine-letter condition were on average merely 6.63% above endorsements for ungrammatical sequences.

A possible explanation for this result is that although materials in both conditions are linguistic, the One-word stimuli conform more with people's expectations of language. Specifically, people's (learnt) knowledge of language includes knowledge of affixes, and the knowledge that prefixes and suffixes can be attached to stems and thereby change the meaning and sometimes the category of a word. For example, the verb *believe* can be turned into an adjective by attaching the suffix *–able*, thus creating *believable*. The meaning of *believable* can be changed by attaching the prefix *un-* to form *unbelievable*. Moreover, in colloquial English, words can have infixes, such as in *abso-flippin'-lutely* (note that the hyphens here are inserted in order to highlight the infix *flippin'*). The final *-lutely* is dependent on the initial *abso-*, yet this regularity is disrupted by the intervening *flippin'*. It is possible that the knowledge of these regularities underlies participants' performance in the One-word condition. By contrast, the Nine-letter condition, and particularly the form the sequences took during the experiment may not immediately conform with what people associate with language: In natural languages, regularities are not found between sequences consisting of single separated segments. It therefore seems reasonable to argue that it is for this reason that the non-adjacent dependencies in the Nine-letter condition were not detected to the same extent as in the One-word condition.

Importantly, the fact that the non-adjacent regularities were reliably detected in the One-word condition, and that performance in the One-word condition is on a level with performance in the –Gestalt condition (as indicated by the fact that the difference scores did not differ in the Tukey post-hoc) shows that obvious interword breaks to highlight the level of analysis are not required for the successful identification of non-adjacencies in visually presented sequences. Interestingly, these findings conflict with Newport & Aslin's (2004) results, as in the present One-word condition participants reliably detected non-adjacencies between syllables of one "word". Although there are a number of differences between the present AGL paradigm and the design employed by Newport and Aslin, as outlined in section 3.1, the conflicting findings are most likely due to variability. Newport and Aslin did not ensure that sufficient variability between adjacent units directed people's attention onto non-adjacent regularities. In keeping with Gómez's (2002) paradigm, the present grammar involved 24 intervening X elements, making transitional probabilities between the first CVC element and the adjacent X item $p = 0.0417$. Newport and Aslin did not reduce transitional probabilities between adjacent units this far. For example, in their Syllable Language of the form AXB (with A and B representing the non-adjacent dependency between two CV syllables) transitional probabilities between A and X were $p = 0.25$. As argued in Chapter 2, it thus seems to be the case that in instances of insufficient variability between adjacent units, the successful detection of non-adjacent regularities relies on additional cues. Newport and Aslin realised this by harnessing the Gestalt Principle of Similarity in their Segment Languages, in which the non-adjacencies were identified despite relatively high transitional probabilities between adjacent elements.

With regard to pause cues, Peña, Bonatti, Nespor and Mehler (2002) showed that these play a significant role in the detection of non-adjacencies. However, as pointed out in the previous chapter, Peña et al.'s materials were flawed as they included Gestalt cues in addition to pause cues. The findings from both Experiment 1 and 2 challenge these findings, and demonstrate that the extraction of non-adjacent dependencies in the linguistic domain is possible even when the relevant units are not made salient by additional cues, such as Gestalt cues or obvious interword breaks.

Mueller, Bahlmann and Friederici (2008) also argue for the importance of pause cues. Yet interestingly, their behavioural results show that the non-adjacent dependencies were learnt even when there were no gaps between sequences. Still, the learning effect was much better (better by 30%) with pauses. However, once again, they included salient phonological cues in their materials. In their design, words that participated in non-adjacent dependencies only contained the vowels /i/ and /e/ whereas words in category X contained /ɔ/ and /u/. Since the test involved making grammaticality judgements on sequences of the form AXB versus AXX, participants could have based their decisions merely on whether the last CVCV word of the sequence contained an light vowel (/i/ or /e/) or a dark vowel (/ɔ/ or /u/). Alternatively, people might be noting the similarity between the first and the third word of each of the sequences, facilitated by the Gestalt principle of Similarity. The

AXX violations then would have violated the principle and thereby assisted people in making the correct discrimination. The question thus remains as to whether they actually tested participants' ability to reliably distinguish grammatical from ungrammatical sequences based on the acquisition of non-adjacent dependencies or whether their participants based their grammaticality judgements on salient Gestalt cues that allowed them to form a mapping between the first and the third elements. Certainly, the latter case cannot be excluded.

Previous research on the detection of non-adjacent dependencies in the linguistic domain to date has focused on the role of pause cues or Gestalt cues. Experiments 1 and 2 provide convincing evidence showing that in visually presented sequences, and with sufficiently low transitional probabilities between adjacent elements, non-adjacent dependencies are reliably extracted without additional cues to highlight the relevant units. However, the Experiments 1 and 2 also demonstrate that people's prior knowledge of language might be of importance when extracting non-adjacencies. Performance levels in the One-word and –Gestalt condition are equal, and, crucially, better than in the Nine-letter condition. The key difference between the three conditions is that non-adjacent dependencies in natural languages are found on the syntax-level (as in the –Gestalt condition) and on the word-level (as in the One-word condition). Yet non-adjacencies are not found between chunks within a sequence of separate segments. It is therefore argued that people's prior expectations of what regularities are found in language may have guided them toward detecting the non-adjacent dependencies in Experiments 1 and 2.

However, a potential confound of the materials used in Experiment 2 has to do with the phonotactics of the English language. Phonotactics play an important role in language learning as it has been shown to be used for speech stream segmentation by infants (Johnson & Jusczyk, 2001; Mattys & Jusczyk, 2001; Mattys, Jusczyk, Luce & Morgan, 1999). In the present context, specifically in the One-word condition, people's knowledge of phonotactically illegal chunks in English may have assisted them in segmenting the sequences into three CVC chunks. So, since the One-word condition involved CVCCVCCVC strings, CC clusters that are not found in English may have been taken to indicate word boundaries. For example, in the sequence *lumfetfip*, the CC cluster *tf* is illegal within words in English, and people may therefore have interpreted this as a word boundary. Segmenting the sequences into CVC chunks consequently could have led participants to extracting non-adjacencies between separate CVC passages much like in the –Gestalt condition. If participants indeed segmented the sequences in the One-word condition, it is interesting that performance levels in the Nine-letter condition indicate that this string segmentation was not possible when inter-grapheme breaks were included. If anything, this confirms the importance of people's linguistic expectations as it demonstrates that phonotactically illegal CC clusters may have been regarded as word boundaries.

3.2 Concluding remarks

In sum, Experiment 2 showed that the acquisition of visually presented non-adjacent dependencies in the linguistic domain is sufficiently robust that it does not require helpful breaks between CVC chunks to indicate the correct level of processing. Moreover, the findings of Experiment 1 support the view that Gestalt cues are also unnecessary for the detection of non-adjacencies. Taken together, our findings thus far illustrate that non-adjacent regularities between linguistic units can be detected even if they are not highlighted. As a next step, Chapter 4 will explore the role that highlighting the relevant elements of analysis plays in the non-linguistic domain.

Chapter 4

The acquisition of non-adjacencies in the non-linguistic domain

The previous two chapters have shown that people can identify non-adjacent relationships between linguistic elements in a visually presented AGL paradigm with and without assistance from additional cues. However, the findings thus far give no indication as to whether the mechanisms involved in the detection of non-adjacencies are also recruited when the stimuli in question are non-linguistic in nature. The aim of Experiments 3 and 4 was therefore to test participants' non-adjacency learning in the non-linguistic domain.

4.1 Experiment 3

The question of whether the mechanisms underlying the human capacity for language are unique to language or whether they form part of a general-purpose learning device is a fundamental one, and different theoretical approaches were discussed in Chapter 1. Chapter 1 also reviewed experimental evidence for powerful learning mechanisms that are available to humans during language acquisition and that also seem to be accessible in the non-linguistic domain, e.g. in stream segmentation (Saffran, Newport & Aslin, 1996; Saffran, Aslin & Newport, 1996; Abla & Okanoya, 2009; Kirkham, Slemmer & Johnson, 2002; Saffran, Johnson, Aslin & Newport, 1999) and ABA pattern learning (Marcus, Vijayan, Rao & Vishton, 1999; Saffran, Pollak, Seibel & Shkolnik, 2007). Some learning mechanisms employed during language learning are therefore also recruited for non-linguistic learning[10].

Recent years have seen a notable increase of research into visual statistical learning. Fiser and Aslin (2002a) used the statistical structure underlying the linguistic sequences in Saffran, Newport & Aslin's (1996) classic experiment for their visual statistical learning tasks. Here, adult participants were shown an animation of a shape moving horizontally from one side of a PC screen to the other. Along its path, it disappeared behind a centrally located occluder, changing into a different shape every time it reappeared. The transformations that the shapes underwent were ordered as strings of specific triplets. So, for example one of the triplets was Shape A – Shape B – Shape C. Whenever Shape A disappeared behind the occluder, the next to appear was Shape B, which in turn changed into Shape C. This series of experiments revealed that adults show robust extraction of the temporally structured sequences. Fiser and Aslin (2002b) also showed that infants are capable of learning subtle spatial relationships between geometrical shapes in

[10] There is a possibility that these mechanisms are in fact unique to language, and that an identical but separate mechanism operates outside the domain of human languages. This point will be discussed in detail in Chapter 7.

multi-object scenes. In their experiment, 9-month old infants displayed high sensitivity toward co-occurrence of shape pairs as well as toward the probability with which individual elements of pairs co-occurred.

Further investigating visual statistical learning, Turk-Browne, Jungé and Scholl (2005) demonstrated the robustness of this mechanism. Across three experiments, Turk-Browne et al. trained and tested adult participants on serially presented triplets. Of crucial importance in this context was the inclusion of non-adjacent regularities in their visual learning task. Participants were exposed to sequences of shape-triplets taken from two separate sets. One set involved green shapes and the other involved red shapes. During training, the shape sequences from both sets were used in a randomly interleaved fashion. So although Red $Shape_1$ predicted Red $Shape_2$, which in turn predicted Red $Shape_3$ (and similarly for the green shapes), participants were exposed to sequences such as Red $Shape_1$ – Green $Shape_1$ – Green $Shape_2$ – Red $Shape_2$ – Green $Shape_3$ – Red $Shape_3$. By interleaving the two separate sets, Turk-Browne et al. tested the acquisition of non-adjacent dependencies in visually presented sequences. Their findings reveal that the shape-triplets participants were informed to attend to (either green or red), albeit for a cover task, were indeed learned despite the intervening shapes. Turk-Browne et al. argue that this is not surprising as a visual statistical learning mechanism that cannot compute dependencies through noise would be of little use during every-day tasks such as driving. However, due to the experimental set-up, in which participants were informed to attend to merely one set of shapes, it is unclear whether participants were in fact learning the non-adjacent dependencies or whether they were merely ignoring the intervening shapes. Importantly, their findings reveal little about people's ability to extract non-adjacent dependencies across domains as they did not directly contrast their non-linguistic experiment with a linguistic equivalent.

As the key issue addressed in this thesis regards the acquisition of non-adjacent dependencies, the main aim of the present experiments is to explore the domain-specificity of non-adjacency learning in a visual AGL paradigm. Although the identification of non-adjacent dependencies has previously been investigated in both the linguistic (Gómez, 2002; Newport & Aslin, 2004) and the non-linguistic domain (Turk-Browne, Jungé & Scholl, 2005), these experiments are hard to compare directly as they differ along a number of dimensions. For example, Turk-Browne et al. employed a cover task, and their non-adjacent regularities were separated by a variable number of intervening elements. This was not the case in either Gómez's or Newport and Aslin's experiments. For this reason, the present two experiments aim to directly contrast the acquisition of non-adjacent dependencies across the linguistic and non-linguistic domain. To this end, the identical experimental set-up as used in Experiments 1 and 2 was employed here, with non-linguistic (Experiment 3), as well as linguistic and non-linguistic (Experiment 4) materials. The first research question to be answered by Experiments 3 and 4 is (1) Are non-adjacent dependencies in the non-linguistic domain equally learnable as in the linguistic domain when presented visually?

With regard to the visual non-linguistic materials used in previous experiments, the question arises whether these stimuli are adequate. To illustrate the point, consider the visual materials used to replicate Saffran, Aslin and Newport's (1996) experiment in the non-linguistic domain. Saffran, Aslin et al.'s linguistic stimuli involve combinatorial reuse of consonants and vowels (e.g., *golabu* and *bidaku* share the plosive *b* and two vowels). The discrete shapes used by Kirkham, Slemmer and Johnson (2002), and by Abla and Okanoya (2009) are non-combinatorial, in that each word in the Saffran, Aslin et al. stimuli corresponds to a geometrical shape-pair or shape-word. These stimuli differ in both shape and colour (Kirkham et al.) or just shape (Abla & Okanoya) from the other shapes. So there is no re-use of smaller sub-elements corresponding to the re-use of segments in *golabu* and *bidaku*, and the non-linguistic materials are therefore less complex than the equivalent linguistic stimuli used by Saffran, Aslin et al. (1996).

An important aspect of language as put forward by Hockett (1960) is duality of patterning. This feature refers to the fact that language involves recombination on two levels. On the phonemic level, phonemes can be combined in different ways to form larger linguistic units: /k/, /æ/ and /t/ can be combined to form the word *cat* and can then be recombined in a different order to create *act*. Thus, the same three elements are re-used to form two separate meaning-carrying units. On the syntactic level, words can be combined and re-combined in different ways to form sentences. Importantly, this duality of patterning highlights the fact that morphological units in natural languages contain internal structure and thus complexity. In the auditory modality the recombined units are phonemes, and in the visual modality they are graphemes. Crucially, the non-linguistic materials used in previous experiments (e.g. by Fiser & Aslin 2002a; Kirkham, Slemmer & Johnson, 2002; Turk-Browne, Jungé & Scholl, 2005) do not replicate this level of internal structure found in language, as their objects lack internal structure and, importantly, systematic re-use of individual components. This issue will be addressed in the present experiments. A further aim of Experiments 3 and 4 therefore is to address this issue by manipulating the internal complexity of the non-linguistic stimuli. Research question (2) therefore is: What role does stimulus simplicity play in the detection of non-adjacent dependencies in the non-linguistic domain.

4.1.1 Method

4.1.1.1 Participants

A total of 88 adult participants were tested, 69 females and 19 males, randomly assigned to one of three conditions. All participants were recruited from Northumbria University campus and the research centre's pool of regular experimental participants. 32 of the 96 participants took part in the Holistic condition, 32 participated in the Componential condition, and 24 in the Shape condition. They were all native speakers of English, and they received either £4.50 or course credit for taking part.

4.1.1.2 Materials

The experiment consists of three conditions differing in materials. The grammar used for all conditions is the same as in Experiments 1 and 2 (see Table 2.1, p. 59), which again is based on Gómez's 2002 grammar (see Table 1.3, p.35). For the Componential condition, the AL used in the –Gestalt condition (Experiment 1) was translated into complex black and white matrix patterns, in which the internal structure of the non-linguistic patterns was matched with the orthography of the artificial words used in the –Gestalt condition. Thus, each grapheme from the AL corresponds directly to a specific sub-component of the Componential patterns. For example, every "l" from the AL was mapped onto a certain sub-pattern within the 10x10 grid, as was every "u" and every "m", creating a non-linguistic equivalent of the artificial word *lum*. Figure 4.1 shows three examples of artificial words and their corresponding componential patterns.

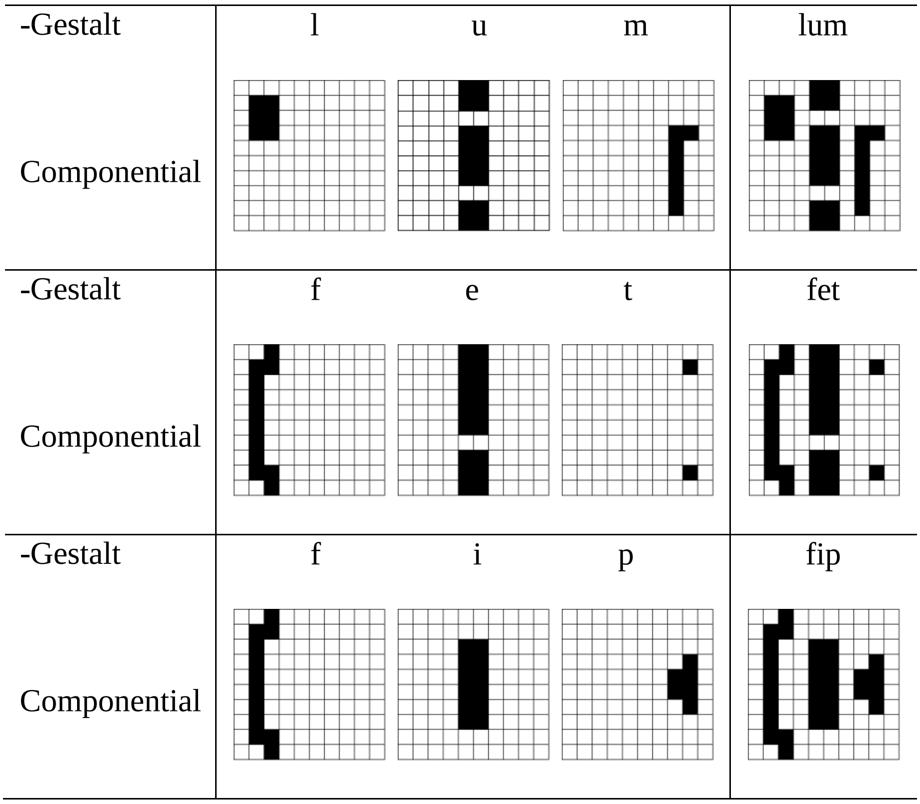


Figure 4.1: Design of the Componential patterns. Each sub-pattern in the Componential condition is mapped onto a grapheme in the –Gestalt condition, which results in each 10x10 black-and-white matrix pattern containing a total of three sub-patterns, much like each artificial word contains three sub-elements (i.e. graphemes).

By re-using sub-patterns to form larger patterns, the Componential stimuli replicate the two layers of patterning that represent a defining feature of natural languages (Hockett, 1960), and thus capture an important linguistic feature not previously used in non-linguistic AGL experiments.

For the Holistic condition, every artificial word from the –Gestalt condition was mapped onto a distinct black and white matrix pattern. Compared with the Componential condition, the stimuli for the Holistic condition were thus generated to appear more like a single unit rather than one pattern containing three sub-components, and crucially, they did not include systematic re-use of sub-components.

The materials for the Shape condition are more in line with materials used in previous non-linguistic sequence learning experiments (e.g. Abla & Okanoya, 2009; Fiser & Aslin, 2002a, Kirkham, Slemmer & Johnson, 2002). Again, each of the non-linguistic stimuli corresponds to a specific word from the –Gestalt condition, and is instantiated as a coloured shape. Thus, each non-linguistic stimulus in the Shape condition differs from the others along two dimensions: shape and colour. Figure 4.2 illustrates how the example string *lum fet fip* was translated into the three non-linguistic conditions. All non-linguistic materials can be found in Appendix A.

Research on visual scene perception is relevant in the present context for two reasons. Firstly, as suggested by Field (1987), visual environments exhibit statistical structure in as much as that pixel intensities in natural scenes are all patterned in a similar fashion. Adjacent pixels tend to have similar intensities, and the more intervening pixels there are, the more different the intensities become. Perception of visual scenes thus involves analysing the underlying statistical structure in a similar way to language. Secondly, high-level vision always involves acquiring information by mapping visual stimuli onto meaning, i.e. identifying relevant objects and scenes (Henderson & Hollingworth, 1999). In this way, visual scene perception is not entirely unlike language acquisition. Visual scene perception entails a combination of saccadic eye movements and fixations, whereby the acquisition of visual information can only take place during fixations (Henderson, 2007) as the anatomy of the human retina only allows for fixated areas to be processed in detail

(Hollingworth & Henderson, 2002). As fixation is then a requirement for information to be processed, a question of primary concern is what areas within visual scenes tend to be fixated. If this is done in a bottom-down fashion, then stimulus properties, for example colour or contrast, will attract the viewer's attention and thus result in fixation. From a top-down perspective, the gaze would be guided by prior knowledge of similar scenes or cognitively controlled due to the task demands (Henderson, 2007). Nuthman and Henderson (2010) propose that visual scene perception is most likely a result of both manners of processing. In their view, people will scan visual scenes and initially focus their attention on objects guided by their knowledge. Longer fixation and more detailed analysis of visual materials, however, is object-based. Longer fixation and consequent acquisition of information is therefore based on whether the initial scene search resulted in the identification of an area, which was then deemed informationally meaningful and thus worth fixating. In Nuthman and Henderson's theory of scene perception, people tend to parse scenes, and subsequently select and analyse relevant objects. In the present AGL experiments, specifically in the Componential condition, correctly parsing the three-element sequences resulting in the identification of the first and third pattern as highly meaningful, would result in acquiring the underlying grammar.

In all three non-linguistic conditions (Experiments 1 and 2), the materials were presented visually, on a white computer screen. As in Experiment 1, each three-element sequence was displayed for 2500 ms, all three elements of each string were presented simultaneously, and each presentation of a string was separated by a 1000 ms display of a blank (i.e. white) screen. Again, the experiment was created using Slide Generator.

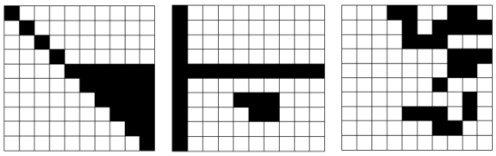
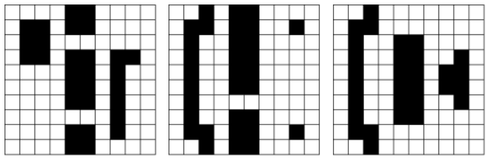
Condition	**Stimuli**		
-Gestalt	lum	fet	fip
Shape			
Holistic			
Componential			

Figure 4.2: Design of all three non-linguistic materials based on the –Gestalt stimuli. In the Shape condition each shape corresponds to a specific artificial word, and in the Holistic condition each matrix pattern is also mapped onto a certain word from the –Gestalt materials. In the Componential condition, the materials are orthographically matched with the CVC words. This design resulted in an increase of internal structure and importantly, with the Componential stimuli being the closest non-linguistic match with the –Gestalt materials.

4.1.1.3 Procedure

The procedure for Experiment 3 was the same as for Experiments 1 and 2. Again, each condition consisted of four subgroups (Language 1a, Language 1b, Language 2a and Language 2b, see Tables 2.1 (p.59) and 2.2 (p.63). Participants in each condition were trained on sequences generated by the PSG containing non-adjacent dependencies, and were required to make grammaticality judgments on individually presented strings in the subsequent testing phase. 50% of all sequences presented on test contained unfamiliar X elements, i.e. stimuli not encountered during training, to test participants’ ability to extract the underlying grammar and apply it to sequence with a novel surface structure (see Table 2.2, p. 63).

Before the training phase participants were told they were going to view a large number of sequences consisting of three shapes (in the Shape condition) or three black and white patterns (in the Holistic and Componential conditions). After training, participants were informed that all sequences followed a specific rule, and that the testing phase required them to decide whether the sequence presented on test followed the same rule or not. As in Experiments 1 and 2, they indicated their response by key press.

4.1.2 Results

Table 4.1 (p.104) reports mean percentage endorsement for grammatical and ungrammatical sequences, containing familiar and novel X items, across all three conditions.

Three separate independent-samples t-test were conducted to compare percentage correct responses between L1 and L2 in the three conditions[11]. Levene’s test for equality of variances was non-significant in all three t-tests ($p > .130$). The t-tests revealed no significant differences between L1 and L2 in any of the three conditions (Shape: $t(22) = .398$, $p = .694$; Holistic: $t(30) = .648$, $p = .522$; Componential: $t(30) = .068$, $p = .946$), and therefore Language (L1 versus L2) was not included as a factor in the subsequent analysis.

[11] The non-linguistic materials for the Holistic and Shape conditions were randomly mapped onto artificial words from the –Gestalt condition from Experiment1, and the sub-patterns of the Componential condition were randomly matched with graphemes used for the –Gestalt condition. Due to this random assignment, the subgroups were not included as a factor in the following analysis as potential interactions would give no insight with regards to domain-specificity.

Table 4.1: Mean percentage endorsements (and standard deviations) for grammatical and ungrammatical strings with familiar and unfamiliar X items, across three conditions.

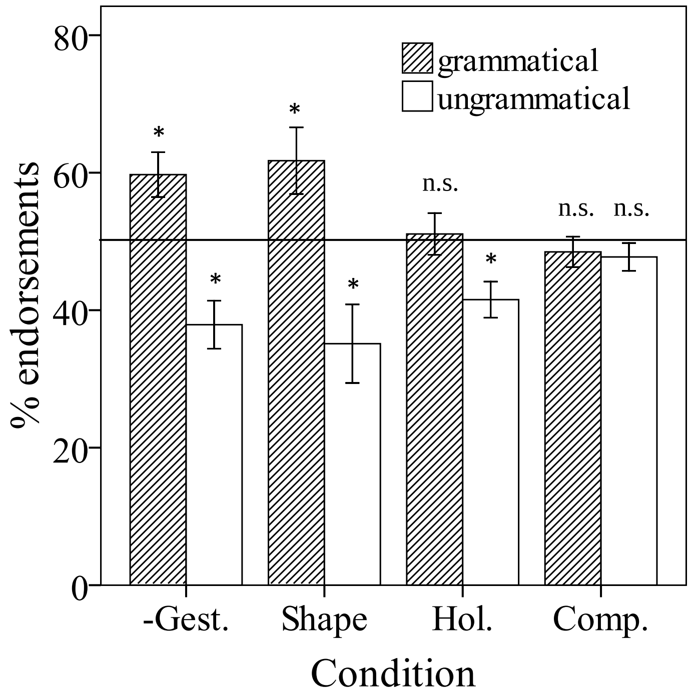
Condition	**Grammaticality**	**Familiarity**	**% endorsements**	**SD**
Shape	grammatical	familiar	77.55	21.37
		novel	45.95	35.55
	ungrammatical	familiar	45.37	34.26
		novel	24.88	28.26
Holistic	grammatical	familiar	65.28	17.92
		novel	36.89	25.46
	ungrammatical	familiar	55.56	21.97
		novel	27.52	17.98
Componential	grammatical	familiar	54.08	13.20
		novel	42.88	15.13
	ungrammatical	familiar	51.91	10.92
		novel	43.58	14.56

Since the key issue of this chapter regards domain-specificity and the ability to track non-adjacencies across the linguistic and non-linguistic domain, the present results were analysed along with the previous results for the –Gestalt condition from Experiment 1. Therefore, an ANOVA on percentage endorsements across all non-linguistic conditions and the –Gestalt condition was carried out, with Condition (Shape, Holistic, Componential, -Gestalt) as a between-subject factor, and Familiarity (familiar X item versus unfamiliar X item) and Grammaticality (grammatical versus ungrammatical) as within-subject factors. The complete results of the ANOVA can be found in Table 4 in Appendix B. The analysis resulted in two main effects and three interactions.

As in Experiments 1 and 2, the present ANOVA revealed a main effect for Familiarity and for Grammaticality. The Familiarity effect, $F(1,116) = 107.43$, $p <$

.001, is due to the fact that across all conditions, people endorsed more with strings containing familiar X elements (57.86%) than with strings containing unfamiliar X elements (37.98%). The main effect for Grammaticality, $F(1,116) = 25.82$, $p < .001$, was due to participants overall favouring grammatical strings (55.26%) over ungrammatical strings (40.57%).

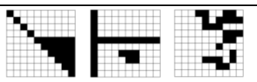
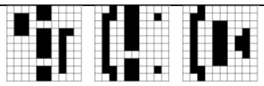
Most importantly, there was a significant Grammaticality x Condition interaction, $F(3,116) = 4.07$, $p = .009$ (see Figure 4.3). To determine whether and to what extent the non-adjacent dependencies were detected across conditions, a series of paired-samples t-tests were carried out to contrast percentage endorsements for grammatical strings against percentage endorsements for ungrammatical sequences. As four of these t-tests were conducted, a corrected alpha-level of $\alpha = .0125$ was employed.



Figures 4.3: Mean percentage endorsements for the Grammaticality x Condition interaction. Horizontal line across indicates chance level (50%). Asterisks indicate performance significantly different from chance level.

Participants' ability to track non-adjacencies in the –Gestalt condition was established in Chapter 2. The present paired samples t-tests indicate that this ability also operates in the non-linguistic domain as people successfully discriminated between grammatical and ungrammatical sequences in the Shape condition (p = .010). For the Holistic and for the Componential conditions, the t-test resulted in a non-significant finding (see Table 4.2). The t-tests therefore clearly show that performances differed across the conditions. Yet to even further unpack the interaction, the difference scores (i.e. percentage endorsements with grammatical – percentage endorsements with ungrammatical sequences) for all four conditions were submitted to a Tukey post-hoc analysis.

Table 4.2: Mean percentage endorsements for grammatical and ungrammatical strings, along with results from paired samples t-tests for the –Gestalt (taken from Experiment 2), Shape, Holistic and Componential (Comp.) conditions.

Condition	Example string	Grammaticality	% endorsements	Statistic
-Gestalt	lum fet fip	grammatical	59.72 (18.46 SD)	t(31) = 3.61,
		ungrammatical	37.89 (19.75 SD)	p = .001
Shape		grammatical	61.75 (23.75 SD)	t(23) = 2.82,
		ungrammatical	35.13 (28.00 SD)	p = .010
Holistic		grammatical	51.09 (17.15 SD)	t(31) = 1.89,
		ungrammatical	41.54 (14.84 SD)	p = .068.
Comp.		grammatical	48.48 (12.52 SD)	t(31) = .40,
		ungrammatical	47.74 (11.44 SD)	p = .695

This analysis confirmed that participants in the Componential condition (.74%, 10.56 SD) were significantly outperformed by people in the –Gestalt (21.83%, 34.26 SD), p = .041, as well as by participants in the Shape condition (26.62%, 46.30 SD), p = .015, with regards to discrimination of strings based on grammaticality. Discrimination between grammatical and ungrammatical strings in the Holistic condition (9.55%, 28.59 SD) was not different from any of the other conditions. There were no other significant pair-wise comparisons.

In line with previous analyses, endorsement rates for grammatical and ungrammatical sequences were also contrasted against chance performance (50%) in separate one-sample t-tests. The t-tests on grammatical sequences showed that participants performed significantly above chance only in the –Gestalt, $t(31) = 2.979$, $p = .006$ (as demonstrated also in Experiments 1 and 2), and in the Shape condition, $t(23) = 2.42$, $p = .024$. The t-tests resulted in non-significance for the Holistic condition, $t(31) = .358$, $p = .723$, and the Componential condition, $t(31) = .686$, $p = .498$. With regard to ungrammatical strings, the one-sample t-tests again showed that participants in the –Gestalt, $t(31) = 3.47$, $p = .002$, and Shape condition, $t(23) = 2.60$, $p = .016$, performed similarly in that they endorsed with significantly less ungrammatical strings than expected by chance. Participants in the Holistic condition also endorsed with significantly less than 50% of all ungrammatical sequences, $t(31) = 3.23$, $p = .003$. In Componential condition, however, the t-test resulted in a non-significant finding, $t(31) = 1.12$, $p = .273$. All these follow-up t-tests therefore show that the Grammaticality x Condition interaction is due to participants extracting the non-adjacencies in the –Gestalt and Shape conditions. In the Holistic conditions, people have had some sensitivity towards the regularities, whereas the underlying grammar was not extracted in the Componential condition.

Moreover, there was a significant Familiarity x Grammaticality interaction, $F(1,116) = 5.72$, $p = .018$ (see Figure 4.4).

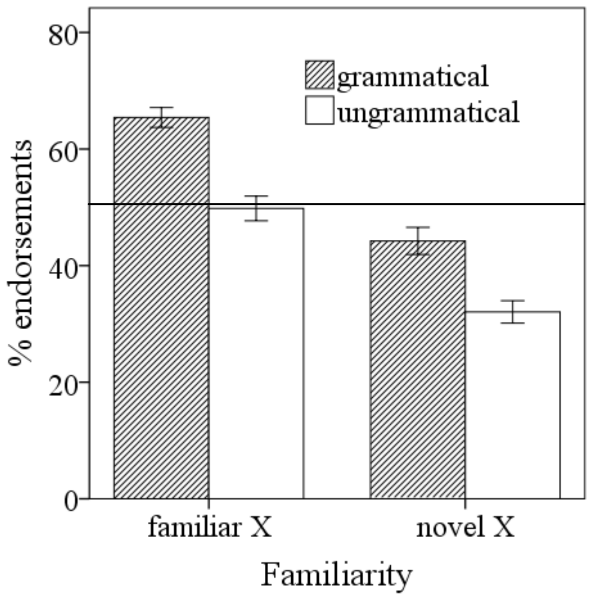


Figure 4.4: Mean percentage endorsements for the Familiarity x Grammaticality interaction, with error bars (+/- 1SE). Black line across indicates chance level performance (50%).

As there were two t-tests, an alpha level of α = .025 was employed. The t-tests show that for strings containing familiar as well as for strings containing unfamiliar X items, participants accepted more grammatical than ungrammatical sequences (see Table 4.3).

Table 4.3: Results from paired samples t-tests for contrasting mean percentage endorsements for grammatical and ungrammatical strings based on familiarity with the X items.

X item	Grammaticality	% endorsements	Statistic
familiar	grammatical	65.42 (18.79 SD)	$t(119) = 5.25, p < .001$
	ungrammatical	49.81 (23.29 SD)	
novel	grammatical	44.24 (25.50 SD)	$t(119) = 3.81, p < .001$
	ungrammatical	32.06 (21.04 SD)	

As these t-tests did not adequately unpack the interaction, the difference scores were calculated (i.e. percentage endorsements for grammatical strings –

percentage endorsements for ungrammatical strings) for strings containing familiar and novel X elements. These were then submitted to a paired-samples t-tests, which revealed a significant difference, t(119) = 2.09, p = .039, between the difference scores for familiar strings (15.60%, 32.58 SD) and unfamiliar strings (12.18%, 35.02 SD). This shows that participants were more likely to correctly discriminate grammatical from ungrammatical strings when they involved familiar X items as opposed to novel X items. However, this reveals nothing about the extraction of non-adjacent dependencies across the different conditions.

The final interaction was Familiarity x Condition interaction, F(3,116) = 5.37, p = .002 (see Figure 4.5).

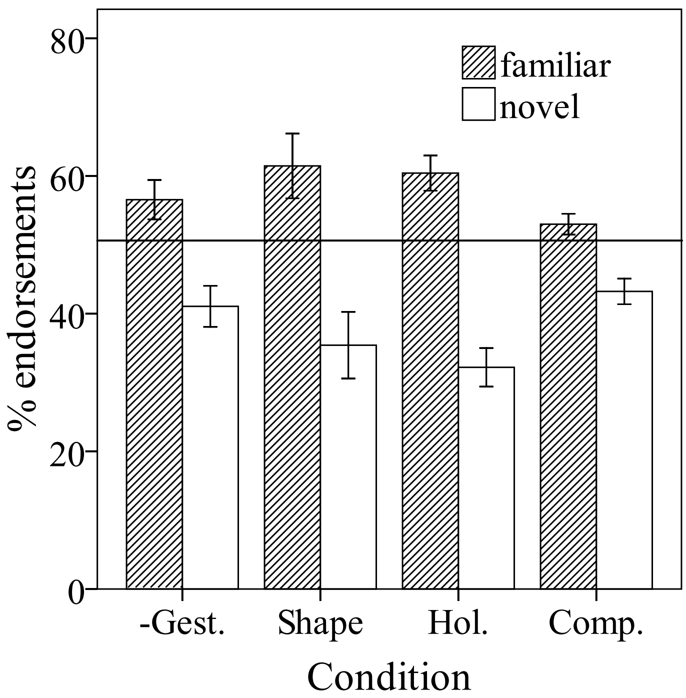
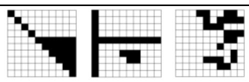
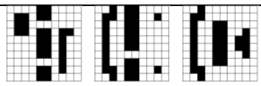


Figure 4.5: Mean percentage endorsements for the Familiarity x Condition interaction, with error bars (+/- 1SE). Black line across indicates chance level performance (50%).

Of key interest with regard to this interaction is whether sequences containing familiar X items were reliably favoured over strings containing novel X items. For this reason, four paired-samples t-tests were conducted (see Table 4.4) to compare percentage endorsements for strings with familiar and unfamiliar surface forms, across all four conditions. Again, an alpha-level of α = .0125 was employed. The t-tests show that in all conditions, strings involving X items encountered during training are significantly favoured over strings containing novel X elements.

Table 4.4: Results from paired samples t-tests for contrasting mean percentage endorsements for familiar and unfamiliar strings for the –Gestalt, Shape, Holistic and Componential (Comp.) condition.

Condition	Example string	Familiarity	% endorsements	Statistic
-Gestalt	lum fet fip	familiar	56.55 (11.34 SD)	t(31) = 5.94,
		unfamiliarity	41.06 (11.16 SD)	p < .001
Shape		familiar	61.46 (17.90 SD)	t(23) = 4.48,
		unfamiliarity	35.42 (19.04 SD)	p < .001
Holistic		familiar	60.42 (14.10 SD)	t(31) = 6.01,
		unfamiliarity	32.20 (16.13 SD)	p <.001
Comp.		familiar	52.99 (9.63 SD)	t(31) = 5.53,
		unfamiliarity	43.23 (13.75 SD)	p < .001

In order to further understand the interaction, the difference scores were calculated by subtracting percentage endorsements for ungrammatical strings from percentage endorsements for grammatical strings. This was done for all four conditions, and the values were subsequently submitted to a Tukey post-hoc test.

This analysis showed that the difference scores in the Componential (9.77%, 14.76 SD) were significantly lower than for the Shape (26.04%, 28.51 SD), p = .023, and Holistic condition (28.21%, 26.57 SD), p = .003. The difference score for the -Gestalt (15.49%, 14.76 SD) was not different from any of the other conditions. None of the other pair-wise comparisons were significant. The most interesting aspect of the Familiarity x Condition interaction was therefore that although sequences containing familiar X items were favoured across all conditions, this effect was particularly strong in the Shape and Holistic conditions.

There were no further main effects or interactions.

4.1.3 Discussion

The purpose of the present experiment was twofold: (1) to investigate domain-specificity in the acquisition of non-adjacent dependencies; (2) to investigate the role of stimulus simplicity in non-linguistic materials. If the detection of non-adjacencies were indeed a completely domain-general ability, participants in all four conditions might have been expected to perform equally well. However, the findings here show that in visually presented sequences, participants only successfully identified non-adjacent dependencies in the Shape condition.

For the non-linguistic conditions, the materials used differed with regard to internal complexity. In the Componential condition, the patterns aimed to replicate the internal complexity of language in the non-linguistic domain. Specifically, language is combinatorial on the phonetic/graphemic level in that it re-uses phonemes/ graphemes to form words. This re-use of elements was employed in the design of the Componential condition, and in this respect the Componential condition is the closest match with the –Gestalt materials. Importantly, participants in the Componential condition were unable to extract the non-adjacent dependencies. A possible explanation for this finding is that participants were misled by the internal structure of the Componential patterns, trying to find regularities within the structures themselves. Importantly, this was not the case in the linguistic domain. In the –Gestalt condition – and indeed in natural language more broadly - people do not get lost in the internal structure of the words. For instance, *lum fet fip* is an example string from the –Gestalt condition. Participants in this condition did not allow the internal structure of the words, i.e. the graphemes, guide their attention away from the regularities that governed the grammaticality of the sequences. So, for example the co-occurrence of the grapheme *f* in initial position of the second and third word in the sequence *lum fet fip* may well have initially struck participants as an important regularity. Crucially, however, the re-use and co-occurrence of specific graphemes was not a sufficiently reliable regularity for people to base their grammaticality judgements on. Instead, in the linguistic domain, people were sensitive to more consistent regularities, which in this case were the non-adjacent dependencies. In contrast to this, it is argued here that there is a possibility that the Componential equivalent sub-pattern of the grapheme *f* , which was repeated in the same way as *f* was repeated in *lum fet fip*, might have misdirected people's attention toward this re-use of the same element. Thus, in the linguistic domain, people seem to know what regularities are important in determining grammaticality, whereas in the non-linguistic domain people do not seem to have these prior expectations. People might be equipped with certain language-specific expectations. In other words: Duality of patterning is indeed a defining feature of language, and – crucially – if is replicated in the non-linguistic domain, people seem to struggle with identifying the relevant units of analysis. Therefore, the prior expectations people bring along when computing language is not automatically carried over into the non-linguistic domain. Importantly, this does not provide evidence for the innateness of these language-specific expectations. Participants here were all native speakers of English. In English, non-adjacent dependencies between segments (i.e. vowels or consonants) within words is not a crucial regularity. Thus, this presumably does not represent part of the present participants' linguistic knowledge. However, this is not the case for all languages. In Finnish, for example, the occurrence of vowels is subject to strict constraints. This is known as Vowel Harmony and dictates that none of the vowels within one word can be from opposing vowel sets. There are three vowel classes in total, namely front vowels /y, ø, æ/ , back vowels /u, o, a/ and neutral vowels /i, e/ (Suomi, McQueen & Cutler, 1997). Therefore, if the initial syllable of a Finnish word contains a front vowel, all following vowels must be either from the class of front vowels or of neutral vowels. Similarly, if the first vowel is a back vowel, then all other vowels must also be either back or neutral. Vowel Harmony found in Hungarian and Turkish operates in a similar fashion (Ohala, 1994). For native speakers of these languages, paying attention to sub-elements (i.e. segments) of words and non-adjacent relationships between them is presumably of high importance. Therefore, the linguistic knowledge displayed by participants in Experiment 3, which can be roughly formulated as a heuristic along the lines of "ignore the internal structure of each of the words as regularities between individual letters do not play an important role" may be learnt knowledge based on the linguistic input received during language acquisition.

Further, this linguistic knowledge that allowed participants in Experiment 1 to learn the underlying structure of the –Gestalt materials most likely feeds into a domain-general learning mechanism. Participants in the Shape condition successfully discriminated grammatical from ungrammatical strings on test, showing that the ability to detect non-adjacencies is not limited to the linguistic domain. In line with the hypothesis presented above to explain the failure of participants to learn the non-adjacent relationships between Componential patterns, it is argued here that participants in the Shape condition succeeded at the task due to the fact that there was no internal complexity to the shapes that could have misdirected people's attention. Thus, eliminating the inner structure of the non-linguistic stimuli and creating simple non-linguistic materials highlighted the correct level of analysis. Specifically, the Shape condition only allows for regularities to operate between individual shapes, as there are no re-used sub-elements which could provide an alternative level of analysis. Additionally, the Shape materials contain redundant cues. Cue redundancy has been the topic of an extensive amount of research, with the main focus being on the processing of audio-visual redundant cues in infancy (Bahrick, Lickliter, Castellanos & Vaillant-Molina, 2010; Lewkowicz, 2000). For example, seeing a ball bounce up and down as well as hearing the ball bounce up and down provides synchronous information about a specific temporally structured sequence across the visual and auditory modality. The ability to integrate synchronous information across two modalities offers a much better understanding of the infants' surroundings, and in early development, infants seem to be more reliant on redundant cues than later on in life (Bahrick & Lickliter, 2000). For the present Shape materials, the redundant cues remain within the visual modality. Nevertheless, the coloured shapes allow participants to encode the shapes as well as the colours, which could both result in successful acquisition of the underlying grammar.

The Holistic condition represents an intermediate stage as the materials contain some internal structure, but no systematic re-use of sub-patterns. Importantly though, these materials were also too complex for participants and the non-adjacent dependencies were therefore not detected.

Finally, Experiment 3, much like Experiments 1 and 2, also identified the importance of surface familiarity for grammaticality judgments. Participants clearly took the familiarity of the X items into consideration when accepting or rejecting sequences on test. The Familiarity x Condition interaction showed that specifically in the Componential condition, participants' familiarity with the X items had a smaller effect on endorsements. This may indicate that the Componential patterns are too complex, and not sufficiently distinguishable from each other. This issue will therefore be addressed in detail in Chapter 6.

With regard to the research questions posed for Experiment 3, the present results suggest that the detection of non-adjacencies is a domain-general ability and therefore operates in both the linguistic and non-linguistic domain, as shown by performance levels in the –Gestalt and Shape conditions. However, stimulus simplicity plays a crucial role. The Shape condition highlights the relevant units of analysis by removing other potential levels of analysis (i.e. the sub-lexical level), thereby assisting people in the extraction of non-adjacencies. Thus, domain-specific expectations might play a role: Adults may be equipped with language-specific knowledge, which guides their attention toward the relevant units in linguistic sequence learning but not in non-linguistic sequence learning.

4.2 Experiment 4

The sequences in Experiment 3 were presented simultaneously as it has been suggested (Saffran, 2002) that this makes the underlying regularities more salient. However, in order to reach a true account of the mechanisms involved in language learning, the manner of presentation must also be addressed. Although simultaneous presentation may remove or at least reduce working memory effects on grammaticality judgements as it does not require participants to store individual elements of each sequence, language itself is not always simultaneously presented. In fact, in the auditory modality language is a sequentially ordered signal. This issue will therefore be addressed in Experiment 4.

Experiments 1 and 3 showed that people can track non-adjacent dependencies in the linguistic domain without assistance from additional cues, and that non-adjacencies in the non-linguistic domain are learnable as long as the stimuli used are simple in appearance. The working hypothesis is that this can be explained by language-specific knowledge, which feeds into a domain-general learning device. It could therefore be expected that the mode of presentation, whether simultaneous or sequential, should not result in a different pattern of findings. At the same time, due to the fact that language is primarily presented sequentially, there is a possibility that the learning mechanism underlying language learning is only available, or more likely to be used, when the materials are presented sequentially.

Thus, Experiment 4 investigates the detection of non-adjacencies by carrying out visual sequence learning experiments on sequentially presented materials. Specifically, the research question here asks whether sequentially presented materials render the same pattern of results as simultaneously presented strings did in

Experiments 1 and 3. For this reason, Experiment 4 used materials from the +Gestalt, -Gestalt, Shape and Componential conditions.

4.2.1 Method

4.2.1.1 Participants

24 adult participants took part in each of four conditions, rendering a total of 96 participants. There were 64 females and 32 males. All participants were native English speakers and received either course credit or £4.50 for taking part. All participants in the –Gestalt condition and 15 participants in the Shape condition were recruited from Aberdeen University's Psychology undergraduate population, the remaining participants were recruited from Northumbria University campus.

4.2.1.2 Materials

The materials used for this experiment were identical to the ones used in Experiments 1 and 3. For the sequential +Gestalt and –Gestalt conditions, the AL from Experiment 1 was used, the +Gestalt condition thus involved the inclusion of the length cue to highlight the non-adjacencies (see Table 2.1, p. 59). The Shape and Componential materials used were identical to the ones from Experiment 3. Experiment 4 did not involve a Holistic condition as the previous experiment showed that the non-adjacent dependencies were not learnt in this condition of intermediate complexity. Importantly then, the present experiment was more focused on the difference in performance levels between non-adjacencies between simple shapes and non-adjacencies between Componential patterns.

In both the linguistic and non-linguistic conditions, the three elements for each sequence were presented one at a time. Before each string, a fixation cross appeared for 500ms. Each element of each string was presented for 833ms. As each string consisted of three units, this rendered 2499ms (3 * 833ms) total exposure time to each string. Exposure time was therefore very close to the 2500ms participants received for each string during simultaneous presentation. There were 200ms pauses (i.e. blank screen) between each element and 700ms pauses between strings.

4.2.1.3 Procedure

As in the previous experiments, this experiment involved a training phase followed by a testing phase. Initially, participants were informed that they were going to be exposed to a large number of sequences made up of three shapes (in the Shape condition), three black and white patterns (in the Componential condition) or three made-up words (in the +Gestalt and –Gestalt conditions), and that although they were only going to see one element on the screen at a time, each sequence was made up of three units. After the training, they were told that all the sequences they had seen followed a specific rule, and that they were required to decide, for each sequence presented during the test, whether that sequence followed the same rule as the training sequences. Again, 50% of all test items contained an unfamiliar X item to avoid participants merely memorising sequences.

Unlike in the previous experiments, this experiment did not include any sub-groups (i.e. versions a and b, see Table 2.2, p. 63) as this factor did not produce significant effects in Experiments 1 and 2. However, to control for potential biases people might have for certain elements, each of the conditions still contained two versions, L1a and L2a as shown in Table 2.1 (p.59).

4.2.2 Results

Table 4.5 (p. 119) shows mean percentage endorsements for grammatical and ungrammatical strings containing familiar and novel X items for all four conditions.

Table 4.5: Mean percentage endorsements (and standard deviations) for grammatical and ungrammatical strings with familiar and unfamiliar X items, across all four conditions.

Condition	**Grammaticality**	**Familiarity**	**% endorsements**	**SD**
Shape	grammatical	familiar	86.81	20.12
		novel	68.52	35.37
	ungrammatical	familiar	28.59	36.24
		novel	13.43	20.31
+Gestalt	grammatical	familiar	73.15	25.15
		novel	67.59	29.87
	ungrammatical	familiar	33.33	31.79
		novel	27.55	26.26
-Gestalt	grammatical	familiar	71.30	21.29
		novel	57.52	29.58
	ungrammatical	familiar	43.63	33.01
		novel	28.70	25.44
Componential	grammatical	familiar	59.60	14.63
		novel	53.13	18.76
	ungrammatical	familiar	51.04	18.68
		novel	45.49	13.15

First, an independent samples t-test was carried out to compare the number of correct responses between Aberdeen students and Northumbria students in the Shape condition in order to ensure there was no difference between the two student populations. It revealed that there was no significant difference in performance, $t(22) = .842$, $p = .20$, and this factor was therefore not included in the following analyses. Next, four separate independent t-tests were conducted to test whether the percentage of correct responses differed between L1a and L2a in each condition.

The results show that in the -Gestalt condition, there was a significant difference. As the Levene's test for equality of means was significant ($p < .001$), equal variances are not assumed, and therefore the corrected values are reported here: $t(17.647) = 2.146$, $p = .046$. This significant difference is due to participants in L2a (72.68%, 26.26 SD) giving a mean of 18.81% more correct responses than participants in L1a (53.88%, 15.23 SD). For this reason, Language (L1a, L2a) was included as a factor in the following analysis.

A 4 x 2 x 2 x 2 ANOVA on percentage endorsements was carried out, with Condition (+Gestalt, -Gestalt, Shape and Componential) and Language (L1a and L2a) as between-subject factors, and Familiarity (familiar and unfamiliar X items) and Grammaticality (grammatical and ungrammatical) as within-subject factors. The complete results for the ANOVA can be found in Table 5 in Appendix B. The ANOVA resulted in two main effects and one interaction.

As for Experiments 1 to 3, the ANOVA resulted in a main effect for Familiarity, $F(1, 88) = 25.84$, $p < .001$, and a main effect for Grammaticality, $F(1, 88) = 56.74$, $p < .001$. The descriptive statistics reveal that participants generally tended to endorse with approximately 10% more familiar strings (55.93%) than strings containing a novel X element (45.24%). Familiarity therefore played a significant role in participants' acceptance of strings on test. The main effect for Grammaticality is due to the fact that, disregarding all other factors, participants were more likely to accept grammatical strings (67.2%) than ungrammatical strings (33.97%).

Importantly, there was a significant Grammaticality x Condition interaction, F(3, 88) = 5.35, p = .002 (see Figure 4.6). There were no further main effects or interactions.

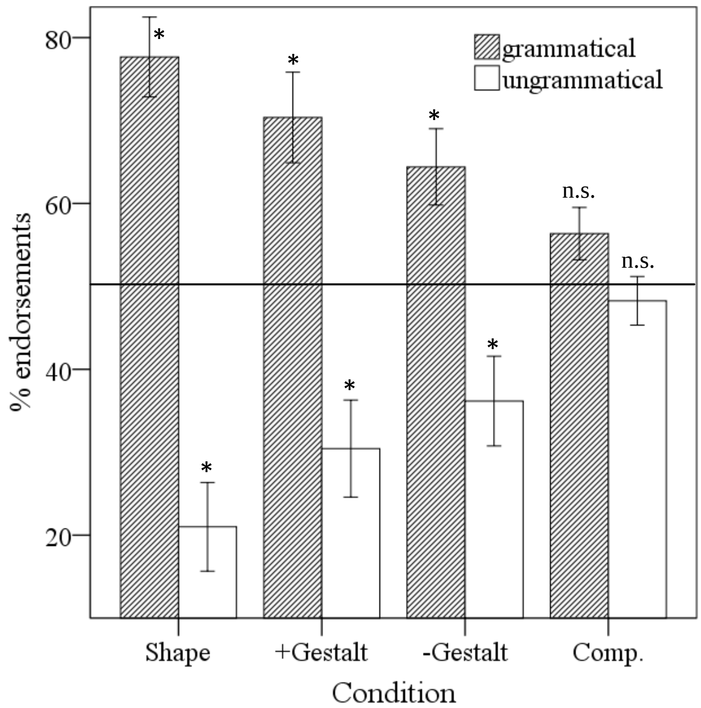
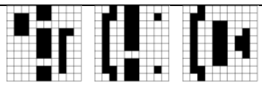


Figure 4.6: Grammaticality x Condition interaction. Mean percentage endorsements for all four conditions, with error bars (+/- 1 SE). The black bar across indicates chance performance (50%). Asterisks indicate performance significantly different from chance.

The interaction was followed up with a series of paired-samples t-tests to compare mean percentage endorsements for grammatical and ungrammatical strings across the four conditions (see Table 4.6). A corrected alpha-level ($\alpha = .0125$) was employed. The results show that grammatical and ungrammatical sequences were reliably discriminated in both linguistic conditions as well as in the Shape condition. In the Componential condition, by contrast, discrimination between sequences based on grammaticality was non-significant, which accounts for the interaction.

Table 4.6: Mean percentage endorsements for grammatical and ungrammatical strings, along with results from paired samples t-tests for the sequential +Gestalt, –Gestalt, Shape and Componential (Comp.) conditions.

Condition	Example string	Grammaticality	% endorsements	Statistic
+Gestalt	lum fetac fip	grammatical	70.37 (26.71 SD)	t(23) = 3.844,
		ungrammatical	30.44 (28.65 SD)	p = .001
-Gestalt	lum fet fip	grammatical	64.41 (22.58 SD)	t(2.98) =
		ungrammatical	36.17 (26.50 SD)	2.98, p = .007
Shape		grammatical	77.66 (23.57 SD)	t(23) = 5.86,
		ungrammatical	21.01 (26.20 SD)	p < .001
Comp.		grammatical	56.37 (15.42 SD)	t(23) = 1.66,
		ungrammatical	48.26 (14.37 SD)	p = .111

To further explore the interaction, the difference scores were calculated (percentage endorsements with grammatical – percentage endorsements with ungrammatical sequences) and submitted to a Tukey post-hoc analysis. The analysis revealed that discrimination between grammatical and ungrammatical sequences in the Shape condition (56.66%, 47.39 SD) was significantly better than in the Componential condition (8.10%, 23.98 SD), p = .001. The difference scores in the +Gestalt (39.93%, 50.89 SD) and –Gestalt (28.24, 46.41 SD) conditions were not significantly different from any of the other conditions. There were no further significant pair-wise comparisons.

As in the previous Experiments, one-sample t-tests were carried out to contrast endorsement rates with grammatical and ungrammatical sequences in all conditions against 50% chance performance. The results show that in the Shape, t(23) = 5.75, p < .001, as well as in both linguistic conditions (+Gestalt: t(23) = 3.74, p = .001; -Gestalt: t(23) = 3.13, p = .005) participants accepted a significantly higher proportion of all grammatical sequences than expected by chance. In the Componential condition, the one-sample t-test resulted in a marginal significance,

$t(23) = 2.02$, $p = .055$. For ungrammatical strings, participants in the +Gestalt, $t(23) = 3.34$, $p = .003$, -Gestalt, $t(23) = 2.56$, $p = .018$, and Shape condition, $t(23) = 5.42$, $p < .001$ accepted significantly less ungrammatical strings than would be expected by chance. In the Componential condition, by contrast, this was not the case, $t(23) = .59$, $p = .56$. This confirms that the grammar was learnt in the linguistic condition and the non-linguistic condition involving simple, unanalysed shapes. In the non-linguistic condition involving componential patterns, the non-adjacent dependencies were not extracted.

4.2.3 Discussion

The aim of this experiment was to investigate the effect of sequential presentation on the detection of non-adjacent dependencies in the visual modality. In general, sequential presentation renders the same pattern of results as previously found in the simultaneous AGL experiments: Non-adjacent dependencies are learnable in both linguistic conditions (+Gestalt and –Gestalt) and in the non-linguistic condition involving simple shapes (Shape condition). However, participants in the Componential condition failed to extract the non-adjacent relationships. Thus, these findings are consistent with the hypothesis of a domain-general learning tool, which integrates language-specific information: Non-adjacencies between linguistic and between simple non-linguistic elements are learnable, whereas non-adjacent dependencies between non-linguistic units that correspond to language with regards to internal complexity cannot be extracted. The re-use of sub-elements in the linguistic domain does not misdirect people’s attention onto irrelevant patterns, whereas in the non-linguistic domain this seems to be the case. However, removing the internal structure of the complex materials eliminates the (spurious) sub-element – to – sub-element level of analysis in the non-linguistic domain: Under these circumstances, non-adjacencies between non-linguistic stimuli are learnable, as demonstrated by successful discrimination between grammatical and ungrammatical sequences in the Shape condition.

However, the question remains whether the mode of presentation matters to the learnability of non-adjacencies based on stimulus type. For this reason, a cross-experimental analysis was carried out, in which endorsement rates for crucial stimulus types were submitted to an ANOVA.

4.3 Cross-experimental analysis

For reasons of clarity, Table 4.7 (p.125) summarises mean percentage endorsements with simultaneously and sequentially presented grammatical and ungrammatical sequences for Shape, +Gestalt, -Gestalt and Componential materials from Experiments 1, 3 and 4.

In order to explore whether the mode of presentation had an effect on the acquisition of non-adjacencies, a 2 x 4 x 2 ANOVA was conducted on percentage endorsements, with Presentation Mode (simultaneous, sequential) and Stimuli Type (+Gestalt, -Gestalt, Shape and Componential) as between-subjects factors and Grammaticality as a within-subjects factor. The full ANOVA results are displayed in Table 6 in Appendix B. The ANOVA resulted in one main effect and two interactions.

Table 4.7: Mean percentage endorsements for grammatical and ungrammatical strings across all four conditions, organised according to mode of presentation.

Stimuli type	Presentation mode	Grammaticality	% endorsements	SD
+Gestalt	simultaneous	grammatical	68.52	21.72
		ungrammatical	37.56	26.23
	sequential	grammatical	70.37	26.71
		ungrammatical	30.44	28.65
-Gestalt	simultaneous	grammatical	59.72	18.46
		ungrammatical	37.89	19.75
	sequential	grammatical	64.41	22.58
		ungrammatical	36.17	26.50
Shape	simultaneous	grammatical	61.75	23.75
		ungrammatical	35.13	28.00
	sequential	grammatical	77.66	23.57
		ungrammatical	21.00	26.19
Componential	simultaneous	grammatical	48.48	12.52
		ungrammatical	47.74	11.44
	sequential	grammatical	56.37	15.42
		ungrammatical	48.26	14.37

The main effect for Grammaticality, $F(1,200) = 97.12$, $p < .001$, was due to more grammatical strings (63.41%) than ungrammatical strings (36.78%) being accepted on test. The two interactions were Grammaticality x Presentation Mode, $F(1,200) = 5.96$, $p = .016$ (see Figure 4.7), and Grammaticality x Stimuli Type, $F(3,200) = 9.27$, $p < .001$. There were no further main effects or interactions.

The Grammaticality x Stimuli Type interaction was entirely expected, as the previous ANOVAs for the simultaneously presented as well as for the sequentially presented materials resulted in a Grammaticality x Condition interaction. The interaction is therefore due to successful discrimination between grammatical and ungrammatical strings for the linguistic materials and for the Shape stimuli, but not for the Componential patterns (see descriptive statistics in Table 4.7, p.125).

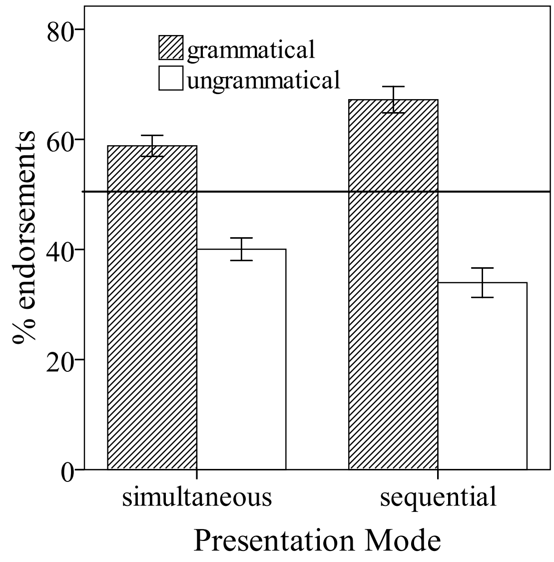


Figure 4.7: Grammaticality x Presentation Mode interaction, with error bars (+/- 1 SE). The black line across indicates chance level performance (50%).

Since the key issue is the discrimination of strings based on grammaticality, the significant Grammaticality x Presentation Mode interaction was further investigated in two paired-samples t-tests, contrasting mean percentage endorsements for grammatical and ungrammatical sequences across both presentation modes (see Table 4.8). A corrected alpha-level was employed (α = .025).

Table 4.8: Results from the paired-samples t-test contrasting mean percentage endorsements for grammatical and ungrammatical based on mode of presentation.

Presentation mode	Grammaticality	% endorsements	Statistic
simultaneous	grammatical	58.83 (20.17 SD)	t(111) = 5.58, p < .001
	ungrammatical	40.04 (21.81 SD)	
sequential	grammatical	67.20 (23.45 SD)	t(95) = 7.03, p < .001
	ungrammatical	33.97 (26.14 SD)	

The t-tests show that in general, grammatical strings were favoured over ungrammatical sequences across both presentation modes. To further explore the interaction, the difference between grammatical and ungrammatical strings was computed separately for both modes of presentation. These difference scores were subsequently contrasted in an independent samples t-test, which resulted in a significant finding, $t(176.84) = 2.54$, $p = .014$ (equal variances not assumed as the Levene's test was significant, $p < .001$). Thus, participants exposed to sequentially (33.23%, 46.34 SD) presented materials were more likely to successfully detect the underlying grammar than participants exposed to simultaneously presented stimuli (18.79%, 35.64 SD).

4.4 General discussion

The cross-experimental analysis compared performances in the acquisition of linguistic and non-linguistic non-adjacent dependencies across two different modes of presentation: simultaneous and sequential. It has been previously argued that in sequence learning experiments, structural regularities are easier to identify when the strings are presented simultaneously (Saffran, 2002). However, the present experiments do not confirm this. The Grammaticality x Presentation Mode interaction demonstrated that presentation mode has a significant effect on participants' grammaticality judgments. Crucially, discrimination between sequences based on grammaticality was better in sequential as opposed to simultaneous presentation.

Language is first and foremost instantiated as speech, and speech is a temporally ordered signal and by consequence sequential. Importantly, non-adjacent dependencies are a defining feature of natural languages (see Chapter 1). A potential explanation for the present finding is therefore that although the materials were presented visually, sequential presentation may have activated the mechanism underlying the detection of non-adjacent dependencies. Sequentially ordered non-adjacencies may thus be easier to detect due to the fact that language itself is sequentially ordered.

Experiments 3 and 4 also provided evidence that the detection of non-adjacent dependencies in the visual modality is – to a degree – a domain-general ability as the regularities between words as well as between shapes were reliably learnt, which is also consistent with the findings of Turk-Browne, Jungé and Scholl (2005), who also found that people are readily capable of detecting regularities between serially presented non-adjacent simple shapes. Although their participants were subjected to the additional cognitive strain of two sets of structurally dependent sequences (red and green) randomly interleaving each other and thus forming non-adjacent dependencies, still they correctly identified approximately 77% of sequences they attended to during training (as opposed to approximately 49% of unattended sequences), which is comparable to the ~77% endorsement with grammatical sequences in the serially presented Shape condition in Experiment 4. People thus seem to be particularly adept at computing structural relationships when unanalysed shapes are used in the non-linguistic domain.

However, the present results suggest that the detection of non-adjacent dependencies is a domain-general ability which is modulated by domain-specific knowledge. This knowledge guides people's attention towards the regularities that matter in language, and crucially, this knowledge is not accessible (or at least not employed) in the non-linguistic domain. These findings show that the issue of domain-specificity requires refined investigations. In Chapter 1, theoretical implications and different approaches to domain-specificity were discussed in detail. Based on the present findings, Atkinson and Wheeler's (2004) view is particularly pertinent. Reconsider Table 4.9.

Table 4.9: In-principle information-mechanism couplings. Taken from Atkinson and Wheeler (2004).

		Information	
		Domain-general	**Domain-specific**
Mechanisms	**Domain-general**	1. Domain-general mechanisms coupled with domain-general information	2. Domain-general mechanisms coupled with domain-specific information
	Domain-specific	3. Domain-specific mechanisms coupled with domain-general information	4. Domain-specific mechanisms coupled with domain-specific information

According to Atkinson and Wheeler, domain-specificity and domain-generality can refer to the stimulus input as well as the mechanisms dealing with this input. They maintain that, in theory, there is no reason to assume that any mechanism, domain-general or domain-specific, can deal only with domain-general or domain-specific information.

The present findings provide experimental support for Atkinson and Wheeler's more detailed approach to domain-specificity. Experiment 3 and 4 showed that the cognitive mechanism underlying the successful identification of non-adjacent dependencies operates in the linguistic (+Gestalt and –Gestalt conditions) as well as in the non-linguistic domain (Shape condition). In accordance with the four options that Atkinson and Wheeler put forward (see Table 4.9, p.129), this narrows potential information-mechanism couplings down to options 1 and 2 – i.e. a domain general mechanism, which might integrate either domain-general (option 1) or domain-specific (option 2) information. However, the non-adjacent regularities were not extracted when they were realised as Componential patterns in the non-linguistic domain. The hypothesis put forward was that of domain-specific knowledge, i.e. knowledge about language, and specifically knowledge about over which units non-adjacent dependencies (can) operate in language. This shows that the information the relevant learning tool draws from is specialised, or at least a subsystem within this learning mechanism seems to draw from domain-specific information. This specialised sub-mechanism can deal with highly complex linguistic stimuli, which exhibit combinatorial re-use of graphemes, yet it cannot analyse non-linguistic materials of comparable complexity. In line with Atkinson and Wheeler, it then seems that the cognitive mechanism underpinning the acquisition of non-adjacent dependencies in the visual modality may itself be domain-general, but it may nevertheless be capable of utilising domain-specific knowledge. This would then represent Atkinson and Wheeler's option 2. Domain-specificity is a key aspect of this thesis, and will therefore be further discussed in Chapter 7.

However, there are a number of further potential explanations for people's inability to track non-adjacent regularities between Componential patterns, which are not related to duality of patterning and domain-specific knowledge. The Componential patterns might be too similar and thus hard to distinguish from each other. This confusability would lead to poor performance during grammaticality judgements. Alternatively, Saffran, Pollak, Seibel & Shkolnik (2007) found that stimulus familiarity influences infant rule learning of non-linguistic stimuli. In line with this finding, the unfamiliarity of participants with the Componential patterns could be the reason for the poor performance levels. These possibilities will be addressed in detail in Chapter 6.

4.5 Concluding remarks

This chapter aimed to investigate domain-specificity in visually presented non-adjacent dependencies and found that the cognitive mechanism involved in the detection of non-adjacent dependencies is domain-general, but that stimulus simplicity plays a crucial role. Only when non-linguistic materials are designed to highlight the relevant units of analysis are the non-adjacencies identifiable. When the internal structure of language is replicated in the non-linguistic domain, the non-adjacent dependencies are not detectable. This is evidence for a domain-general tool which accesses domain-specific knowledge. Furthermore, the mode of presentation matters. When the materials are presented sequentially, people are more likely to accept grammatical sequences. This may be due to the fact that language is first and foremost an auditory signal, and consequently sequentially ordered. The following chapter will therefore explore the acquisition of non-adjacent dependencies in an aurally presented AL.

Chapter 5

Modality effects on the acquisition of non-adjacent dependencies

The previous chapter investigated the acquisition of non-adjacent dependencies in the visual modality. However, language is, first and foremost, an aurally transmitted signal. To this end, this chapter will explore the acquisition of linguistic and non-linguistic non-adjacencies in the auditory modality.

5.1 Experiment 5

There has been much research into modality constraints on how people process and perceive information. Of specific importance to sequence learning are modality effects on serial recall. The recency effect refers to people's ability to better recall the most recent, i.e. the final, element of a sequentially ordered list, and it has been shown to be strongest in the auditory modality (Glenberg & Fernandez, 1988) compared with other modalities. By contrast, Beaman (2002) has demonstrated that the opposite, being able to better recall the initial element of a list, is particularly strong in the visual modality. Moreover, modality constrained learning effects have also been investigated within AGL paradigms (Conway & Christiansen, 2006), with research showing that a stronger recency effect in the auditory and a stronger primacy effect in the visual modality also exists when the sequences are generated by a FSG (Conway & Christiansen, 2009). Investigating sequence learning across three modalities, Conway and Christiansen (2005) found a number of different learning effects between the tactile, visual and auditory modality. Participants in their design were trained and tested on strings generated by a FSG, which were implemented in the tactile modality by vibrotactile pulses to specific fingers, in the visual modality by the sequentially ordered appearance of black squares on a computer screen, and in the auditory modality by strings of musical notes. Conway and Christiansen group the findings into quantitative and qualitative effects. Quantitatively, they found that sequence learning was best in the auditory modality compared with the tactile and visual modality. Qualitatively, the main finding was that, much like in serial recall tasks, there is an advantage of the auditory modality with regards to recency effects.

These definite cross-modal variations show that each of the senses may be attuned to different features of a sequential input. This may mean that the findings of Experiments 1 – 4, and specifically the hypothesis regarding the role of language-specific expectations in the detection of non-adjacencies, could be bound to the visual modality. In order to address this, the present experiment's aim is to explore non-adjacency learning in the auditory modality.

Much research on non-adjacent dependencies has focused on the auditory modality, both in the linguistic (Gómez, 2002; Newport & Aslin, 2004; Peña, Bonatti, Nespor & Mehler, 2002) and non-linguistic (Creel, Newport & Aslin, 2004; Gebhart, Newport & Aslin, 2009; Kuhn & Dienes, 2005) domain. In 2004, Creel, Newport and Aslin investigated the acquisition of non-adjacent dependencies in tone sequences across four experiments. In their design, non-adjacent dependencies were instantiated as musical tones organised as triplets. For example, the triplet F_4_G_4_D_4 represented a non-adjacent dependency to be detected as F_4 reliably predicted G_4, and G_4 reliably predicted D_4. There were two sets of triplets ("odd-numbered" and "even-numbered", p.1122). During training, each triplet from one set was interleaved with a tone-triplet from the other set (here indicated by the underscores). Each tone was presented for 200ms to create a consecutive stream of the form $A_X B_Y C_X D_Y E_X F_Y$…, with the subscripts indicating the two separate sets the triplets were taken from. The testing phases involved a 2AFC, in which one non-adjacent triplet was contrasted with a different non-adjacent triplet that appeared during the training but was scrambled for the test. For example, during their first experiment, the triplet F_4_G_4_D_4 was contrasted against *C#$_4$_G#$_4$_B_4 (which appeared in its grammatical form G#$_4$_C#$_4$_B_4 during the training phase). Creel et al.'s findings suggest that non-adjacent dependencies between tones can only be detected when the relevant tones exhibit featural similarity, which in their design was either pitch or timbre. In their experiment 2, for example, the triplets in one set were high in pitch and the triplets in the other set were low in pitch. The sequences thus had the form $A_{HIGH} B_{LOW} C_{HIGH} D_{LOW} E_{HIGH} F_{LOW}$. Creel et al. argue for the Gestalt Principle of Similarity underlying the detection of non-adjacencies as the perceptual cues (pitch or timbre) allow participants to group the relevant units of analysis more easily.

However, it has been shown that people have strong expectations regarding musical tones (Pearce & Wiggins, 2006) and the patterns they follow (Schellenberg, Adachi, Purdy & McKinnon, 2002). There is therefore a chance that these expectations might have interfered with participants' performances in the Creel et al. experiment. Moreover, similar to what was discussed as the issue of stimulus simplicity in Chapter 4, musical tones are also inadequate equivalents to linguistic materials as they are not sufficiently complex. Previous research has been undertaken to address the issue of stimulus adequacy in the auditory modality. Specifically, Wade and Holt (2005) designed a video game, which participants were trained on for 30 minutes. It involved a number of different alien characters, which all required specific responses on behalf of the participants (e.g. evil aliens had to be shot, good aliens had to be captured). Crucially, each visual occurrence of a certain alien correlated with a specific auditory stimulus. Correct mapping of sounds onto aliens resulted in a better score during the video game as, in higher levels, the auditory sounds preceded the appearance of the aliens on the screen. This, assuming the participant correctly learnt the correlation between sound and alien, allowed the participant to prepare for the right reaction (e.g. shoot or capture) in good time. A similar experiment was conducted by Leech, Holt, Devlin and Dick (2009), who set out to investigate the role of expertise in a non-linguistic categorisation task similar to Wade and Holt (2005) using an fMRI design. Here, training on the video game lasted for a minimum of five hours, and participants' temporal lobe activation during the game was measured prior to training and again after training using fMRI. The sounds used in this experiment were highly complex as they consisted of two spectral peaks layered on top of each other. In doing this, Leech et al. aimed to replicate the complexity inherent in aurally presented phonemes, and they found that neural activation in speech sensitive areas of the temporal lobe post training differed significantly from pre-training. This demonstrates that experience of or expertise with non-linguistic materials drives neural activity in (supposedly) language-specific brain regions, which can be interpreted as evidence for domain-general cognitive capacities. However, Leech et al.'s paradigm involved a categorisation task, which is fundamentally different from the present AGL experiments.

Gebhart, Newport & Aslin (2009) improved this design by using non-musical sounds to investigate the acquisition of non-adjacent regularities. In their experimental design, there were two non-adjacent frames AXB and CXD, and there were two elements in category X, which resulted in transitional probabilities A → X and C → X equalling 0.5, whereas transitional probabilities between non-adjacent sounds (A 9→ B and C→D) remained reliably at 1.0. The question to be investigated was whether participants would be able to segment a continuous stream of sounds based on these statistics. In order to examine this, participants on test were required to discriminate triplets (AXB or CXD) from part-triplets (XDA or BCX). Based on their findings, Gebhart et al. argue that non-adjacencies between non-linguistic elements are only learnable when the relevant elements are perceptually similar, as their participants only successfully discriminated triplets from part-triplets when additional cues rendered the dependencies more salient. Specifically, in their third experiment, the non-adjacencies operated between two similar sounds (raspy) skipping a perceptually different sound (tonal). Only when this cue was included were their participants capable of stream segmentation. Gebhart et al. point toward Newport and Aslin's (2004) findings of non-adjacencies in the linguistic domain only being learnable between segments but not syllables, and therefore argue for the importance of similarity cues (or what was termed Gestalt cues in Chapter 2) in the identification of non-adjacencies. However, although Gebhart et al. as well as Newport and Aslin have strong evidence for the Gestalt principle of similarity being a domain-general mechanism, operating in the linguistic and non-linguistic domain, it does not clearly show that it is a requirement for the successful detection of non-adjacencies. Firstly, Chapter 2 demonstrated that Gestalt cues are not necessary for the acquisition of non-adjacent dependencies. Secondly, Gebhart et al. investigated non-linguistic stream segmentation rather than the extraction of non-adjacent dependencies. Instead of showing sensitivity toward non-adjacent dependencies, participants in their cued experiment could merely have learnt a *raspy-tonal-raspy* pattern, which can be generated by a simple FSG and would have resulted in the same findings. Moreover, the reason why participants in their uncued experiments did not exhibit the ability to segment the stream may well be related to the rather limited number of X elements. Gebhart et al.'s design merely involved two X sounds. In line with Gómez's (2002) variability hypothesis, transitional probabilities between A and X (and between C and X) of 0.5 are much too high to disrupt adjacent regularities and guide the learner's attention toward the non-adjacent dependencies.

Thus, the aim of the present experiment was to remedy the issues of (1) using musical tones (as in Creel, Newport & Aslin), (2) employing stream segmentation to investigate the extraction of non-adjacencies and (3) having high transitional probabilities between adjacent elements (Gebhart et al.). For this reason, Experiment 6 uses the same design and the same underlying grammar as Experiments 1 – 5 to investigate domain specificity in the acquisition of non-adjacent dependencies between sounds differing in timbre. Since domain-specificity remains a key issue for this thesis, the present experiment will also explore the learnability of non-adjacent dependencies between words from the ALs used in Experiment 1. Specifically, the research question is whether the same pattern of results regarding domain-specificity found in the visual modality will hold for the auditory modality.

5.1.1 Method

5.1.1.1 Participants

The participants were Northumbria University students who took part for either course credit or £4.50. They were all native English speakers, 20 males and 76 females, and they were randomly allocated to one of four conditions. Therefore 24 participants were tested per condition.

5.1.1.2 Materials

The experiment consisted of four conditions aimed to replicate the +Gestalt, -Gestalt, Shape and Componential conditions from previous experiments in the auditory modality. The grammar used for this experiment was thus the same one as used previously, which can be found in Table 2.1 (p. 59). In order to minimise the possibility of including unwanted prosodic cues, the materials for the linguistic conditions (which can be found in Table 2.2, p. 63) were generated using synthesised speech. Specifically, the Victoria voice on the Apple Mac OS X built-in speech synthesiser was employed to create the stimuli for the +Gestalt and –Gestalt conditions. Each of the monosyllabic artificial words was between approximately 410ms and 580ms long, whereas the bisyllabic words tended to be longer (540 - 876ms).

The stimuli for the non-linguistic conditions were generated using Logic®[12]. For the auditory Shape[13] condition, each shape from the visual Shape condition was mapped onto a specific 600ms sound. For the auditory Componential condition, each matrix pattern from the visual Componential condition was translated into sounds. Each of the Componential stimuli was also 600ms long in total, but was composed of 3 different 200ms sounds. Therefore, each sub-pattern corresponded to a specific sound, and these sounds were combined to create the Componential stimuli. Crucially, all sounds used in this experiment differed in timbre and not in pitch. Thus, they were all C_3 (middle C) notes instantiated as different instrumental plug-ins. The exact materials (phonetic transcriptions and notes) used for this experiment can be found in Appendix C[14].

In all conditions, a fixation cross appeared for 500ms to indicate the start of a sequence. Each unit, i.e. a word for the linguistic conditions, an unanalysed sound for the auditory Shape condition and a componential sound for the Componential condition, was separated from the next unit of the same sequence by a 200ms pause. There was a pause of 700ms between the end of one string and the start of the following string. As in previous experiments, each of the present sequences lasted for approximately 2500ms.

[12] http://www.apple.com/uk/logicstudio/

[13] Note that the auditory Shape condition was termed as such as the materials used for this condition were the closest match to the visual Shape condition (this will be discussed in 7.2). As this experiment was conducted in the auditory modality, individual elements here were obviously instantiated as sounds rather than shapes.

[14]Additionally, the sound files for all conditions in the present experiment can be found on my private webpage: http://sites.google.com/site/jenniferasturm/

5.1.1.3 Procedure

The procedure used in previous experiments was adapted for the present experiment. Thus, participants took part in a training phase followed by a testing phase. Before the training, they were informed that they were going to hear a large number of sequences consisting of three made-up words (in the +Gestalt and –Gestalt conditions), three sounds (in the Shape condition) or three sounds, which were themselves made up of three sounds (in the Componential condition). There were two opportunities for participants to take breaks during the training phase, which usually lasted for 20 – 25 minutes. Once participants had completed the training, they were told that all sequences they had been exposed to followed a specific rule, and that the testing phase required them to make a decision for individual sequences on whether or not the string followed this rule. Pressing “Y” indicated a yes-response and “N” indicated a no-response. Like in previous experiments, the V and B keys of the keyboards were marked as “Y” and “N”, counterbalanced across participants.

To control for potential biases people might have for certain units, each of the conditions contained two versions, L1a and L2a as shown in Table 2.2 (p. 63). In both versions, each of the 24 X elements appeared with each of the non-adjacent dependencies three times, rendering a total of 216 (3 repetitions x 3 dependencies x 24 X units) sequences during training. On test, participants had to make grammaticality judgments on a total 144 sequences, as each of the three dependencies appeared 24 times in its grammatical form and 24 times containing a grammatical violation. The ungrammatical strings used can be found in Table 2.2 (p. 63). Moreover, 50% of all test items contained an unfamiliar X item to test whether participants would generalize the grammar to sequences with novel surface forms.

5.1.2 Results

Table 5.1 shows mean percentage endorsements for grammatical and ungrammatical sequences containing familiar and novel X elements across the four auditory conditions.

Table 5.1: Mean percentage endorsements (and standard deviations) for grammatical and ungrammatical strings, containing familiar and novel X items, across all four conditions.

Condition	Grammaticality	Familiarity	% endorsements	SD
+Gestalt	grammatical	familiar	71.64	15.12
		novel	50.23	24.71
	ungrammatical	familiar	52.08	25.25
		novel	33.22	17.97
-Gestalt	grammatical	familiar	66.90	25.1
		novel	58.80	26.26
	ungrammatical	familiar	40.39	28.94
		novel	29.28	21.14
Shape	grammatical	familiar	66.09	17.47
		novel	55.56	16.46
	ungrammatical	familiar	55.79	16.09
		novel	48.84	15.97
Componential	grammatical	familiar	58.22	14.34
		novel	57.87	12.06
	ungrammatical	familiar	56.83	15.52
		novel	54.98	14.09

First, four separate independent-samples t-tests were conducted to contrast percentage correct between L1a and L2a in the four conditions. The analysis showed that performance differed significantly in the Componential condition, $t(22) = 2.348$, $p = .028$, due to participants in L1a responding correctly (52.78%, 3.92 SD) more than participants in L2a (49.48%, 2.8 SD). For this reason, Language (L1 and L2) was included as a factor in the subsequent analysis.

Percentage endorsements in all conditions were submitted to a 4 (+Gestalt, -Gestalt, Shape, Componential) x 2 (L1 vs L2) x 2 (familiar vs unfamiliar X elements) x 2 (grammatical vs ungrammatical strings) ANOVA. The full results of the ANOVA can be found in Table 7 in Appendix B. The analysis resulted in four main effects and two interactions.

The main effect for Language, $F(1,88) = 4.96$, $p = .029$, is due to there being a higher percentage of endorsements in L2a (55.87%) than in L1a (51.22%). Although the mean difference of approximately 4.66% is significant, this finding gives no insight into the detection of the non-adjacencies as this factor was not involved in an interaction with Grammaticality. Further, there is no obvious reason why this effect goes in the same direction across linguistic and non-linguistic materials, as these are only arbitrarily linked.

The main effect for Condition, $F(3,88) = 3.51$, $p = .020$, reflects the fact that overall percentage endorsements differed across conditions. A subsequent Tukey post-hoc contrasting percentage endorsements across conditions, revealed that participants in the –Gestalt condition accepted a total of 48.84% of all sequences, which is significantly less ($p = .036$) than participants in the Componential condition (56.97%). The difference between overall percentage endorsement between the -Gestalt and Shape condition was marginally significant ($p = .051$), with participants in the Shape condition accepting a total of 56.57% of all strings. In the +Gestalt condition, participants' endorsements were at 51.79%, which was not significantly different from any of the other percentages. Again, this main effect for Condition gives no indication as to whether participants successfully discriminated grammatical from ungrammatical sequences.

There was a main effect for Familiarity, $F(1,88) = 44.94$, $p < .001$, which is due to the fact that participants, regardless of all other factors, endorsed more with strings containing familiar X units (58.49%) as opposed to novel X items (48.60%). Thus, as in the previous experiments, participants were more likely to accept strings with familiar surface forms.

As in previous experiments, the main effect for Grammaticality, $F(1,88) = 21.87$, $p < .001$, was due to participants overall accepting significantly more grammatical (60.66%) strings than ungrammatical (46.43%) sequences. Importantly, there was a significant Grammaticality x Condition interaction, $F(3,88) = 3.46$ $p = .02$ (see Figure 5.1). In order to further investigate this interaction, and to determine whether the discrimination between grammatical and ungrammatical strings was significant, a series of follow-up t-tests was conducted.

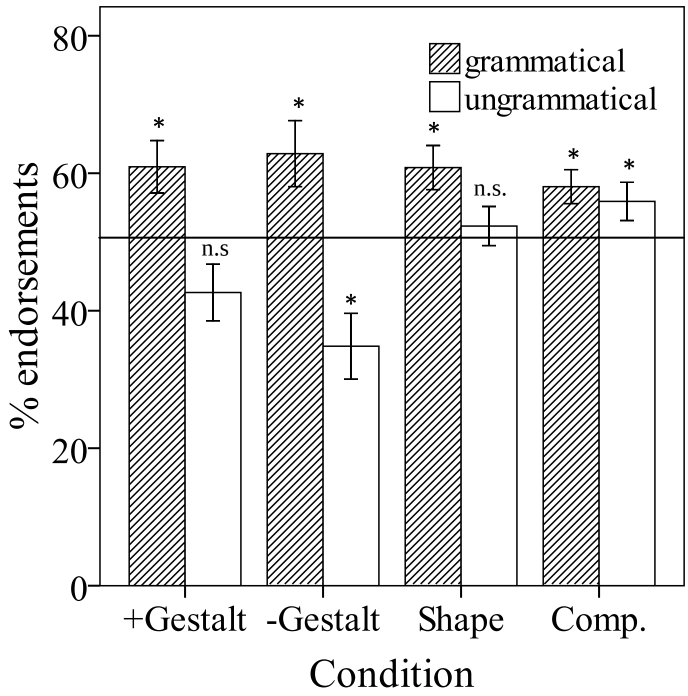


Figure 5.1: Percentage endorsements for the Grammaticality x Condition interaction, with error bars (+/- 1SE). The horizontal line indicates chance level (50%). Asterisks indicate performances significantly different from chance.

First, paired-sample t-tests were carried out to explore whether participants successfully discriminated between grammatical and ungrammatical sequences in each of the conditions (see Table 5.2). A corrected alpha-level was employed (α = .0125).

Table 5.2: Mean percentage endorsements for grammatical and ungrammatical strings, along with results from the paired samples t-tests, for sequences across all four conditions.

Condition	**Grammaticality**	**% endorsements**	**Statistic**
+Gestalt	grammatical	60.94 (18.67 SD)	t(23) = 2.51, p = .020
	ungrammatical	42.65 (20.22 SD)	
-Gestalt	grammatical	62.85 (23.53 SD)	t(23) = 3.21, p = .004
	ungrammatical	34.84 (23.44 SD)	
Shape	grammatical	60.82 (15.75 SD)	t(23) = 2.29, p = .032
	ungrammatical	52.31 (13.95 SD)	
Componential	grammatical	58.04 (12.10 SD)	t(23) = 1.31, p = .204
	ungrammatical	55.90 (13.65 SD)	

The t-tests show that discrimination between grammatical and ungrammatical sequences was only significant in the –Gestalt condition. With the corrected alpha-level, discrimination between sequences based on grammaticality in the +Gestalt as well as in the Shape condition is non-significant. A follow-up Tukey post-hoc analysis was conducted to compare the difference scores (i.e. percentage endorsements with grammatical strings – percentage endorsements with ungrammatical strings) across the four conditions. This analysis showed that the difference scores in the –Gestalt (28.01%, 42.70 SD) were significantly higher than in the Componential condition (2.14%, 8.03 SD), p = .016. The difference scores in the +Gestalt (18.28%, 35.69 SD) as well as in the Shape (8.51%, 18.21 SD) condition do not differ significantly from any of the conditions.

Furthermore, four one-sample t-tests were conducted on endorsement rates with grammatical strings against 50% chance performance. These resulted in a significant finding across all four conditions: +Gestalt ($t(23) = 2.87$, $p = .009$), -Gestalt ($t(23) = 2.68$, $p = .014$, Shape ($t(23) = 3.37$, $p = .003$) and Componential ($t(23) = 3.26$, $p = .003$). Additional one-sample t-tests were conducted to contrast percentage endorsements for ungrammatical strings against chance performance. In the +Gestalt condition, participants did not perform significantly different from chance, $t(23) = 1.78$, $p = .088$. In the –Gestalt condition, however, participants accepted significantly less ungrammatical strings than would be expected by chance, $t(23) = 3.17$, $p = .004$. In the Componential condition, the analysis revealed that participants accepted more ungrammatical strings than would be expected by chance, $t(23) = 2.119$, $p = .045$. Thus, participants in the Componential condition generally accepted more than 50% of all strings, regardless of grammaticality. In the Shape condition, the t-test was non-significant, $t(23) = .813$, $p = .425$.

The Grammaticality x Condition interaction therefore reflects the fact that the non-adjacent dependencies were reliably detected only in the –Gestalt condition. In +Gestalt and in the Shape conditions, people seem to have been sensitive to the regularities to an extent but failed to reliably discriminate between grammatical and ungrammatical sequences on test. In the Componential condition, however, there was no evidence of sensitivity to the non-adjacencies and certainly no evidence for the underlying grammar being detected.

Furthermore, there was a significant Familiarity x Condition interaction, $F(3,88) = 7.03$, $p < .001$ (see Figure 5.2). This interaction was further unpacked with a number of paired-samples t-tests (see Table 5.3, p. 147). The paired-samples t-tests, along with the corrected alpha level of $\alpha = .0125$, showed that including a novel X item on test had no effect on people's endorsements in the –Gestalt and Componential conditions. In the +Gestalt and in the Shape condition, sequences with familiar surface forms were preferred over sequences involving novel X items.

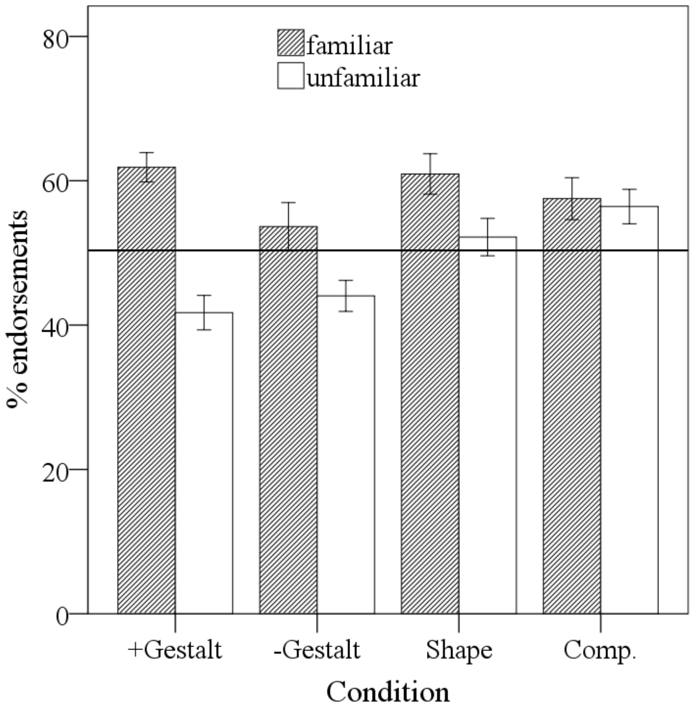


Figure 5.2: Percentage endorsements for the Familiarity x Condition interaction, with error bars (+/- 1SE). Familiarity had little effect on percentage endorsements in the Componential condition.

A subsequent Tukey post-hoc analysis comparing the difference scores (i.e. percentage endorsements with strings containing familiar X item - percentage endorsements with strings containing unfamiliar X item) for each condition showed that the difference between endorsements with familiar and unfamiliar sequences in the +Gestalt (20.14%, 15.25 SD) condition was significantly larger than in the Shape (8.74%, 12.09 SD), $p = .036$, and Componential condition (1.10%, 8.83 SD), $p < .001$. All other pairwise comparisons were non-significant, demonstrating that the difference scores in the –Gestalt condition (9.96%, 19.22 SD) did not differ from any of the other conditions.

Table 5.3: Mean percentage endorsements for familiar and unfamiliar strings, along with results from the paired samples t-tests, for sequences across all four conditions.

Condition	Familiarity	% endorsements	Statistic
+Gestalt	familiar	61.86 (9.99 SD)	t(23) = 6.47, p <.001
	unfamiliar	41.72 (11.71 SD)	
-Gestalt	familiar	53.65 (16.28 SD)	t(23) = 2.45, p = .022
	unfamiliar	44.04 (10.54 SD)	
Shape	familiar	60.94 (13.75 SD)	t(23) = 3.54, p = .002
	unfamiliar	52.20 (12.70 SD)	
Componential	familiar	57.52 (14.23 SD)	t(23) = .61, p = .548
	unfamiliar	56.42 (11.71 SD)	

The Familiarity x Condition interaction is thus due to the familiarity with X items having a significant effect on sequence endorsements in the +Gestalt condition. Familiarity with X elements had a much weaker effect on people's acceptance rates in the Shape condition, and no effect in the –Gestalt and Componential conditions.

There were no further main effects or interactions.

5.1.2.1 Cross-modal comparisons

It has been argued that there are definite modality effects, and specifically that sequence learning in the auditory modality results in superior performance compared with the visual modality (Conway & Christiansen, 2005). For this reason an additional ANOVA was carried out to compare performance between the present Experiment 5 and the results from Experiment 4, in which visual strings were presented sequentially. The materials used for Experiment 5 were as closely matched as possible with the stimuli used for Experiment 4. The full results can be found in Table 8 in Appendix B. Mean percentage endorsements for grammatical and ungrammatical sequences across the visual and auditory modality for all four conditions can be found in Table 5.4.

Table 5.4: Mean percentage endorsements (and standard deviations) for grammatical and ungrammatical strings, for the auditory and visual modality, across all four conditions.

Condition	Modality	Grammaticality	% endorsements	SD
+Gestalt	auditory	grammatical	60.94	18.67
		ungrammatical	42.65	20.22
	visual	grammatical	70.37	26.01
		ungrammatical	30.44	28.65
-Gestalt	auditory	grammatical	62.85	23.53
		ungrammatical	34.84	23.44
	visual	grammatical	64.41	22.58
		ungrammatical	36.17	26.50
Shape	auditory	grammatical	60.82	15.75
		ungrammatical	52.31	13.95
	visual	grammatical	77.66	23.57
		ungrammatical	21.01	26.20
Componential	auditory	grammatical	58.04	12.10
		ungrammatical	55.90	13.65
	visual	grammatical	56.37	15.42
		ungrammatical	48.26	14.37

The 2 x 4 x 2 ANOVA, with Modality (auditory and visual), Condition (+Gestalt, -Gestalt, Shape and Componential) and Grammaticality (grammatical, ungrammatical) as factors, resulted in two main effects and three interactions. There was a main effect for Grammaticality, $F(1,184) = 78.25$, $p < .011$ due to the fact that grammatical sequences were accepted more (63.93%) than ungrammatical sequences (40.20%). The was also a main effect of Modality, $F(1,184) = 4.37$, $p = .038$, as percentage endorsements were higher in the auditory modality (53.55%) than in the visual modality (50.59%). However, neither of these main effects reveal anything about the detection of non-adjacencies across modalities.

The first interaction, Grammaticality x Condition, F(3,184) = 5.47, p = .001, was expected as the previous ANOVAs for Experiment 4 and Experiment 5 showed that the underlying grammar was detected in the +Gestalt, -Gestalt and Shape conditions when materials were presented visually, and were reliably identified in the -Gestalt condition when presented aurally.

The more interesting interactions are the significant Grammaticality x Modality interaction, F(1,184) = 12.53, p = .001 (see Figure 5.3), and Grammaticality x Condition x Modality, F(3,184) = 3.99, p = .009.

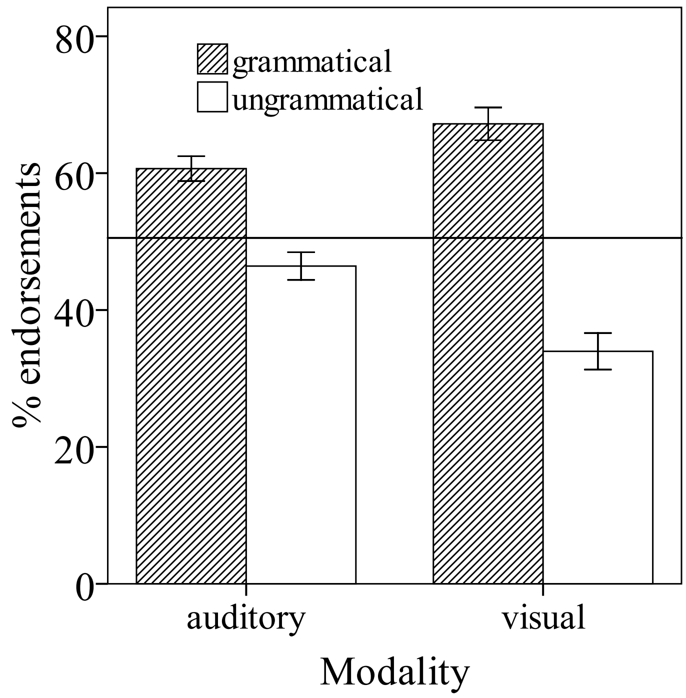


Figure 5.3: Grammaticality x Modality interaction. Mean percentage endorsements for the auditory and visual modality, with error bars (+/- 1 SE). The horizontal line indicates chance level performance (50%).

The Grammaticality x Modality interaction was examined first by conducting paired-samples t-tests to contrast endorsements for grammatical and ungrammatical sequences across the visual and auditory modality (see Table 5.5).

Table 5.5: Mean percentage endorsements for grammatical and ungrammatical strings, along with results from the paired samples t-tests, for sequences across both modalities.

Modality	Grammaticality	% endorsements	Statistic
visual	grammatical	67.20 (23.45 SD)	$t(95) = 7.03$, $p < .001$
	ungrammatical	33.97 (26.14 SD)	
auditory	grammatical	60.66 (17.80 SD)	$t(95) = 4.54$, $p < .001$
	ungrammatical	46.43 (19.83 SD)	

Since these within-modality comparisons failed to reveal the source of the interaction (participants in both modalities successfully differentiated grammatical from ungrammatical sequences), the interaction was further investigated by contrasting the difference scores (i.e. percentage endorsements for grammatical sequences – percentage endorsements for ungrammatical sequences) for the visual and auditory modality in an independent-samples t-test. The Levene's test for equality of variances was violated ($p < .001$) and therefore the corrected values are reported. The t-test showed that discrimination between strings in the visual modality (33.23%, 46.34 SD) was better than in the auditory modality (14.24%, 30.70 SD), $t(164.95) = 3.35$, $p = .001$.

In order to further illuminate the Grammaticality x Condition x Modality interaction, a number of individual 2 x 2 ANOVAs were conducted. The full ANOVA tables can be found in Appendix B (Tables 9 - 12). First a 2 (Modality) x 2 (Grammaticality) ANOVA was carried out to compare percentage endorsements between sequentially presented non-adjacencies in the +Gestalt materials in the visual modality (from Experiment 4) and in the auditory modality (from Experiment 5). The same was done for the –Gestalt stimuli type. For the +Gestalt materials, the ANOVA resulted in a main effect for Grammaticality, $F(1,46) = 21.05$, $p < .001$, as participants endorsed significantly more grammatical (65.65%, 23.29 SD) than ungrammatical (36.55%, 25.30 SD) sequences, regardless of the modality of the input. Importantly, there was no significant Grammaticality x Modality interaction, $F(1,46) = 2.91$, $p = .10$. There were no other significant effects or interactions. The same pattern of findings was revealed for the –Gestalt condition, with a significant effect for Grammaticality, $F(1,46) = 19.09$, $p < .001$. This effect was due to higher acceptance of grammatical (63.63%, 22.83 SD) as opposed to ungrammatical (35.50%, 24.76 SD) strings, regardless of modality. Here, there were also no further significant effects or interactions. This indicates that in the linguistic domain, there are no superior auditory effects when it comes to extracting non-adjacent dependencies based on a PSG.

Next, a 2 (Modality) x 2 (Grammaticality) ANOVA was carried out to compare percentage endorsements between sequentially presented non-adjacencies in the visual Shape condition (from Experiment 4) and auditory Shape condition (from Experiment 5). The analysis resulted in a main effect for Modality, $F(1,46) = 6.34$, $p = .015$, and Grammaticality, $F(1,46) = 39.55$, $p < .001$, as well as in a significant Modality x Grammaticality interaction, $F(1,46) = 21.59$, $p < .001$ (see Figure 5.4).

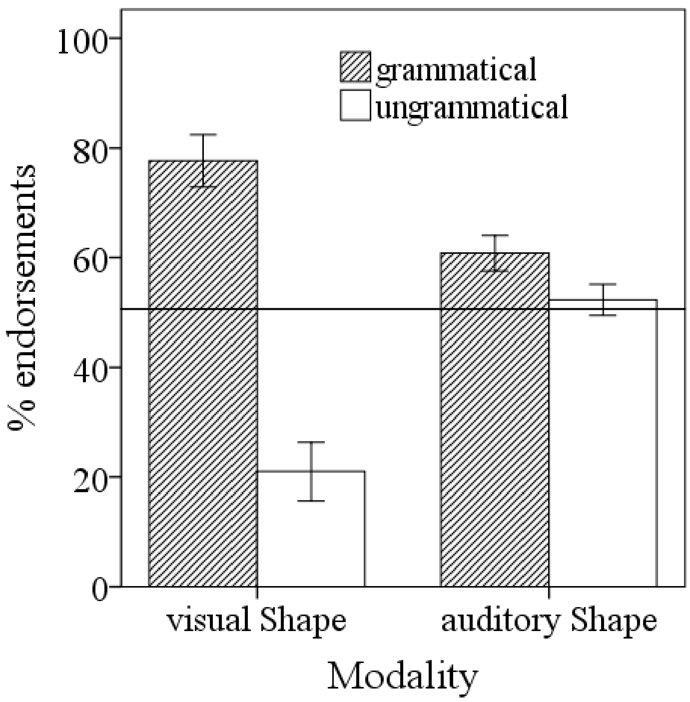


Figure 5.4: Percentage endorsements for the Modality x Grammaticality interaction, with error bars (+/- 1SE). The horizontal line across indicates chance level performance (50%).

Percentage endorsements were higher overall for the aurally presented sequences, with participants here accepting 56.57% of all sequences and thus approximately 7.24% more than participants in the visual-sequential task, who were at 49.33%. This explains the main effect for Modality, yet it reveals little about the acquisition of non-adjacencies. The main effect for Grammaticality is due to a higher percentage of acceptance for grammatical than for ungrammatical strings (69.24% vs. 36.66%), however, this gives little insight into the modality differences. Importantly, the interaction is due to participants in the visual-sequential task being significantly better than participants in the auditory condition at discriminating grammatical from ungrammatical sequences. This was shown in an independent-samples t-test contrasting the difference scores (i.e. percentage endorsements with grammatical – percentage endorsements with ungrammatical sequences) for the

Shape conditions across the two modalities. Levene's test was significant ($p < .001$) and therefore the corrected values are reported. The t-test showed that in the visual modality, discrimination between grammatical and ungrammatical sequences was significantly better (56.66%, 47.36 SD) than in the auditory modality (8.51%, 18.21 SD), $t(29.65) = 4.65$, $p < .001$. As illustrated in Figure 5.4, in the visual modality, participants accepted approximately 77.66% (23.57 SD) of all grammatical and merely 21.01% (26.20 SD) of all ungrammatical sequences, whereas in the auditory task overall percentage endorsements for grammatical strings was 60.82% (15.75 SD) and thus 8.5% higher than for ungrammatical strings (52.32%, 13.95 SD). In fact, discrimination in the visual modality was more than six times better than in the auditory modality. This difference accounts for the interaction.

Finally, the same 2x2 ANOVA was conducted on the Componential materials, contrasting performances in the visual and auditory modality. Here, the main effect for Grammaticality was marginally significant, $F(1,46) = 3.94$, $p = .053$ due to percentage endorsements being slightly higher for grammatical (57.20%, 13.74 SD) than for ungrammatical strings (52.08%, 14.39 SD). This may indicate that some participants might have shown sensitivity toward the regularities. Importantly, however, even if there was some sensitivity toward the non-adjacencies, the fact that the main effect was marginally significant demonstrates that the detection of the underlying grammar instantiated as Componential patterns is much harder. This ANOVA resulted in no further main effects or interactions.

The cross-modal analysis found no further main effects or interactions. In general, the results suggest that modality only had a significant on the Shape materials, and performance levels for all other stimuli types remained at the same level across modalities.

5.1.3 Discussion

The main aim of this experiment was to investigate whether the extraction of non-adjacent dependencies in the auditory modality followed the same pattern as in the visual modality with regards to domain-specificity. The results suggest that this was, in general, the case.

For the linguistic domain, however, the results do not paint a clear picture. The paired-samples t-test (see Table 5.2, p. 144) showed that in the +Gestalt condition, discrimination between grammatical and ungrammatical sequences in the auditory modality did not differ significantly when a strict correction for multiple comparisons was applied to the alpha-level. More importantly, however, the cross-modal analysis revealed no significant Modality x Grammaticality interaction when contrasting performances in the two visual linguistic conditions and the two auditory linguistic conditions. This therefore suggests that adults' ability to extract non-adjacent dependencies in the linguistic domain nevertheless operates equivalently across both modalities. Importantly, much like in visually presented sequences, people in the aural tasks were also reliably capable of tracking these regularities without assistance from the Gestalt Principle of Similarity. In fact, participants in the auditory –Gestalt condition reliably extracted the non-adjacencies (see Figure 5.1, p.143), and, interestingly, familiarity with the intervening X words had no significant effect on endorsement rates in this condition (see Table 5.3, p. 147). This therefore demonstrates that when the underlying grammar is successfully acquired, generalisations to novel forms readily occur.

The underlying grammar governing the sequential input was not detected in the auditory Componential condition. In line with the visual materials, the Componential auditory stimuli were matched element-by-element with the auditory –Gestalt condition and therefore represented the closest non-linguistic equivalent. These findings are in keeping with the results from the previous chapter and are again consistent with the notion of language-specific expectations (not accessible in the non-linguistic domain) that direct learners toward the crucial regularities. Importantly, this domain-specific knowledge cannot be viewed as a fully encapsulated modular mechanism. The non-adjacent regularities between simple, unanalysed sounds were also detected in the auditory Shape condition (although not quite as readily as in the visual modality, which will be discussed shortly). Thus the general ability to track non-adjacent dependencies must be a domain-general ability, which taps into a domain-specific subsystem, i.e. language-specific knowledge.

In addition to the findings regarding domain-specificity, the present experiment also challenges current accounts of modality-constrained sequence learning effects. Conway and Christiansen (2005) strongly argue for superior sequence learning performance in the auditory modality as participants in their auditory task outperformed participants in the visual task by 75% to 62% in the grammaticality judgments. However, in their auditory condition, they used different musical notes. Thus, each stimulus differed from the other along two dimensions: position within the temporal stream and quality of the note. However, in their visual task, participants were exposed to black squares that appeared in a specific order in different locations on the screen. Importantly, every one stimulus in this condition was identical to the next and differed in terms of location and temporal order. The black squares were not qualitatively different from each other in the way the musical notes differed from each other. In contrast to this, the materials used for Experiments 4 and 5 were qualitatively distinct i.e. they looked different and sounded different. For this reason, specifically the visual and auditory Shape conditions of Experiments 4 and 5 might lend themselves more to cross-modal investigations than the materials used by Conway and Christiansen.

However, the cross-modal analysis in 5.1.2.1 revealed that the non-adjacent dependencies in the aurally presented Shape condition were not as reliably detected as in the serially presented visual task, as shown in Figure 5.4 (p. 152). This demonstrates that modality-constrained learning effects require a refined approach and that auditory sequence learning does not necessarily result in better performance as argued by Conway and Christiansen. A potential explanation for Experiment 5's Shape participants' relatively poor performance is that the auditory stimuli used, although qualitatively distinct from one another, were not as distinguishable as the visual materials. The visual materials used in Experiment 4 were different with regards to shape and colour. There is a possibility that this allowed participants to not just encode non-adjacent regularities between shapes but also between colours (or both shape and colour). In the auditory modality, by contrast, participants in the Shape condition were required to extract the dependencies between sounds merely differing in timbre. In a similar way as Conway and Christiansen found better performance in aural sequence learning due to participants being able to extract more distinguishing features, there is a possibility that the cross-modal analysis between

Experiment 4 and 5 shows that participants here performed better in the visual modality due to there being more informative cues available.

It makes sense to assume that learners would extract all available cues in learning tasks. In a related line of research, infants have exhibited sensitivity to a number of different inputs, which they can use to make sense of the world. Frank, Slemmer, Marcus and Johnson (2009), for example, demonstrated that 5-month old infants only succeed at learning an abstract ABA/ ABB rule when instantiated as visually presented shapes and coordinated speech sounds, which shows that infants rely on two modalities, and also information from two domains for successful rule learning. Furthermore, research on spatio-temporal learning has also illustrated infants' ability to integrate multiple, perceptually different, cues (Kirkham, Slemmer, Richardson & Johnson, 2007). Across a total of four experiments, the authors investigated 11-month, 8-month and 5-month old infants' sensitivity towards statistically structured sequences of visual materials. They found that the youngest infants were sensitive to a combination of shape and colour cues but not to the sequential order of the stimuli. The oldest infants, by contrast, showed sensitivity only to the underlying statistics governing the sequences. Importantly, the 8-month old infants demonstrated the ability to extract the underlying statistical input of the sequences only when additional cues of colour and shape assisted them in this task. Kirkham et al. thus suggest that the statistics underlying non-linguistic spatio-temporal sequences may be a cognitively demanding task and, and may not be informative to infants without additional cues until the age of approximately 11 months.

In a similar way, the detection of non-adjacent dependencies has been proposed to be a cognitively demanding task in the linguistic domain (Gómez & Maye, 2005; Newport & Aslin, 2004). Although specifically Experiments 1 and 2 showed that non-adjacencies are learnable by adults without assistance from additional cues, Experiments 3 and 4 clearly demonstrated that in the non-linguistic domain, stimulus simplicity is a crucial factor in the detection of non-adjacencies. Taking all this evidence into account, a potential explanation for the discrepancy found between performance levels in the Shape condition of visually presented shapes (Experiment 4) and aurally presented sounds (Experiment 5) is that the extraction of non-adjacent regularities in the non-linguistic domain is a complex task. As such, it requires additional cues to guide people toward the relevant units. In the visual Shape condition, participants could have encoded the non-adjacent grammar, the colour of the shapes, the shapes or any combination of the three factors. In the equivalent auditory condition, by contrast, participants had less informative cues available to them: the underlying grammar and the timbre.

In sum, the present findings suggest that the detection of non-adjacent dependencies in the auditory modality is subject to the same constraints regarding domain-specificity as in the visual modality. The identification of non-adjacencies in the linguistic domain comes easily to participants regardless of the modality of the input as indicated by the non-significant Modality x Grammaticality interaction. In the non-linguistic domain, however, participants rely on additional guidance in order to track the correct regularities successfully.

5.2 Concluding remarks

The present chapter investigated modality effects on the acquisition of non-adjacent dependencies, and found that the pattern of results in the auditory modality replicates the pattern of results in the visual modality. Thus, in both the visual and auditory modality, people are unable of tracking non-adjacent dependencies when the non-linguistic materials are designed to replicate the internal structure of language, as was done in the Componential materials. Yet non-adjacencies that are instantiated as linguistic elements and simple non-linguistic elements (shapes or sounds differing in timbre) are detected. For this reason, the following chapter’s aim is to further investigate the underlying reason for the poor performances in the Componential conditions. In doing this, the following chapter will focus on the visual modality.

Chapter 6

The role of familiarity in the acquisition of non-adjacencies between complex patterns

Experiments 1 to 5 have established that the extraction of non-adjacent dependencies is, to an extent, specific to language. In the linguistic domain, non-adjacent dependencies are reliably detected, regardless of the perceptual modality and even after the removal of additional cues such as Gestalt cues and breaks. By contrast, non-adjacent dependencies in the non-linguistic domain rely on stimulus simplicity. Experiments 3 to 5 showed that across both the visual and auditory modality, non-adjacencies between complex non-linguistic materials are not learnable when the materials replicate the combinatorial re-use of graphemes/ phonemes found in language. This chapter will further investigate these results by focusing on the role stimulus familiarity plays in the detection of non-adjacent dependencies between componential patterns.

6.1 Experiment 6

There are a number of possible explanations for why the non-adjacent dependencies in Experiments 3 and 4 are not detected when the underlying grammar is instantiated as Componential black and white matrix patterns. In Chapter 5, the case was made for duality of patterning. Language is componential on the graphemic/phonemic level as well as on the higher morphosyntactic level, meaning that graphemes/ phonemes are combined to create morphemes or words, and these are then combined to form sentences. This crucial aspect of language was replicated in the non-linguistic domain by designing the Componential materials, which in both the visual (Experiment 3 and 4) and auditory (Experiment 5) modality, were matched sub-element by sub-element with the artificial language used for the –Gestalt stimuli throughout the work presented in this thesis. Importantly, it is in the Componential conditions that participants failed to extract the non-adjacent dependencies, regardless of the mode of presentation (simultaneous, serial) and modality (visual, auditory). It was argued that language-specific expectations guide people's attention toward relevant units in the linguistic domain, yet fail to do so in the non-linguistic domain. Moreover, it was argued that this linguistic knowledge feeds into a domain-general mechanism capable of tracking non-adjacent dependencies, as the underlying grammar was learnt in the non-linguistic domain as long as the stimuli were simple and unanalysed (i.e. the Shape conditions).

However, there are other potential reasons to consider in explaining people's inability to extract non-adjacencies between Componential patterns. One reason has to do with the matrix patterns themselves. The re-use of sub-patterns for the Componential materials results in complex stimuli. There is thus a possibility that participants struggle to differentiate between individual patterns, resulting in the non-adjacencies not being detected. The fact that Experiment 3 showed that including novel X patterns in the testing phase has much less of an effect on people's acceptance rate[15] may well indicate that they are not even aware of the novel X patterns. The novel items may thus be as unfamiliar to them as the X units encountered during the training. Most importantly therefore, participants' unfamiliarity with the Componential materials may have hindered the extraction of non-adjacent dependencies.

In a related line of research, Saffran, Pollak, Seibel and Shkolnik (2007) examined infant rule learning in the non-linguistic domain across three experiments. It has been found that although infants are adept at abstract rule learning in the linguistic domain (Marcus, Vijayan, Bandi Rao & Vishton, 1999), this ability does not operate equally well in the non-linguistic domain when the rule is instantiated as looming shapes (Johnson, Fernandes, Frank, Kirkham, Marcus, Rabagliati & Slemmer, 2008) or non-linguistic auditory materials such as musical tones and timbres (Marcus, Fernandes & Johnson, 2007). Saffran et al. claim that materials used for non-linguistic rule-like learning in infants frequently exhibited low ecological validity. Specifically, they argue that the difference between performance in the linguistic and non-linguistic domain is due to familiarity. The non-linguistic materials used in previous rule learning studies are entirely novel to infants and are thus not as familiar and indeed meaningful to them as the linguistic materials. Although the linguistic materials used are artificial words and syllables, infants are nonetheless exposed to language on a daily basis and are thus familiar with a linguistic input. The novel materials are, so they maintain, harder to represent and

[15]The Tukey post-hoc analysis showed that the difference between percentage endorsements with strings containing familiar X items and strings containing novel X items was significantly lower than for the other two non-linguistic conditions (Holistic and Shape), see p.106.

more difficult for infants to store in memory during rule learning tasks. Saffran et al. argue that unfamiliar stimuli may be encoded as tokens, i.e. as the specific elements used in the trials. Since individual tokens are not meaningful to the infants, they fail to place the tokens into categories, which in turn means they fail to encode patterns of types. To make their point clearer: If an infant encodes a sequence of shapes, for example triangle-circle-triangle, as a sequence of tokens, he/she will be unable to extract the underlying pattern of types, i.e. CategoryA-CategoryB-CategoryA, and thereby be incapable of applying this ABA rule to novel instantiations of the same rule. Saffran et al. addressed this problem by replicating the Marcus, Vijayan, Rao and Vishton (1999) experiments using non-linguistic stimuli that are familiar to infants, namely pictures of dogs and cats.

In Saffran et al.'s series of experiments, 7-month old infants were habituated to either an ABA or ABB rule using pictures of dogs (Experiment 1) or cats (Experiment 3), or an AAB or ABB rule using pictures of dogs (Experiment 2). On test, infants who were exposed to the ABA rule in Experiment 1 or 3, for example, viewed triads that followed this rule (50%) and triads that followed the ABB rule (also 50%). Across all three experiments, the authors found a novelty effect, with infants looking significantly longer at triads that followed the novel pattern. They conclude that abstract rule learning is a domain-general ability that operates across the linguistic and non-linguistic domain, and that stimulus familiarity is a crucial factor. They argue that using non-linguistic materials that are familiar and therefore meaningful to the infants allowed their infant participants to detect the underlying rule and abstract the rule to new instances on test.

In the present context, the unfamiliarity with the Componential matrix patterns may play an important role in the detection of non-adjacencies. In accordance with Saffran, Pollak, Seibel and Shkolnik's's findings (2007), it seems reasonable to assume that participants' unfamiliarity with Componential patterns may have hindered the identification of non-adjacencies in the AGL task . The aim of the present experiment was therefore to explore whether prior familiarisation with the Componential black and white matrix patterns would result in better performance in the detection of non-adjacent dependencies.

Importantly, very closely related to this issue of familiarity is the issue of confusability. In the visual Shape conditions, for example, the non-linguistic stimuli were also unfamiliar to the participants in as much as they did not merely involve simple shapes, such as rectangles and circles, but the majority of the shapes were fairly abstract (see Table 1 in Appendix A for all the materials used). However, a crucial difference between the Componential materials and the Shape materials was that the shapes – due to the lack of internal, systematic structure – were less confusable and therefore easier to distinguish. By addressing familiarity of the Componential patterns, the present experiment also aimed at targeting the issue of confusability.

In the present experiment, participants were therefore familiarised with the Componential patterns prior to completing the AGL task (from Experiment 3).The amount of familiarisation was manipulated in order to investigate whether the degree of familiarity with the matrix patterns affects the acquisition of the non-adjacent dependencies. The research questions for this experiment therefore were: (1) Does training on the vocabulary of the Componential materials result in higher familiarity with individual patterns? (2) Does increased familiarity with individual patterns result in the detection of non-adjacent dependencies in the AGL experiment?

6.1.1 Method

6.1.1.1 Participants

48 adult participants (eight males and forty females,) took part in one of three conditions, 16 participants in each of the 1-block and 4-block conditions, and 16 in the 8+2-block condition. All participants were native English speakers and they received either course credit or payment for participating. Depending on the condition, participants received either £4.50 (for the 1-block and 4-block conditions) or £10.50 (for the 8+2 condition).

6.1.1.2 Materials

The stimuli for this experiment were the Componential black and white matrix patterns used in Experiments 3 and 4.

6.1.1.3 Procedure

This experiment consisted of two major parts, the familiarisation task and the grammar learning task. There were three conditions due to the between-participants manipulation of familiarisation blocks. In the 1-block condition, participants received one block of familiarisation before moving on to the grammar learning task. The 4-block condition involved 4 blocks of familiarisation prior to the grammar learning task. The 8+2-block condition took place over two days. On Day 1, participants received 8 familiarisation blocks, and on Day two they received a further 2 blocks of familiarisation and consequently progressed to the grammar learning task.

The familiarisation task was conducted on a PC and designed using Experiment Builder (see Figure 6.1 for a schematic illustration of the training and testing phase of the familiarisation task).

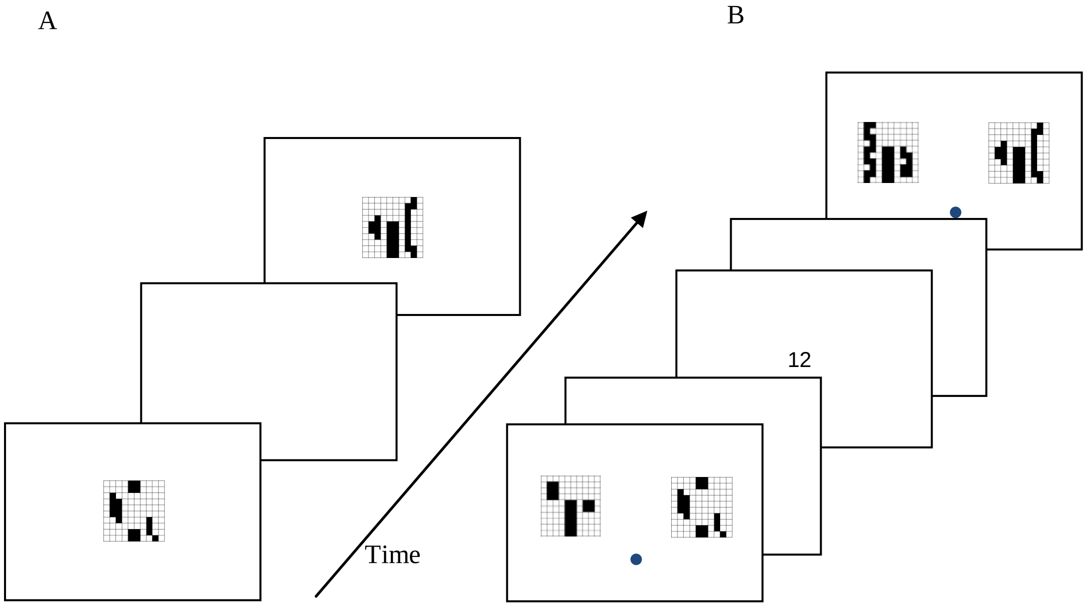


Figure 6.1: Schematic illustration of the familiarisation task. A. Individual presentation of matrix patterns during the training phase. B. 2 AFC task, in which participants were required to click on the pattern they had seen during the training.

Participants viewed the entire vocabulary of the Componential language one item at a time, where a single exposure to a single matrix pattern lasted for 1500ms. Each pattern presentation was separated from the next by a blank screen (250ms). There were 30 patterns in total, made up of the 24 X elements and 6 elements involved in the three non-adjacent dependencies (see Table 1 in Appendix A for all the materials).

During each familiarisation block, each of the 30 patterns was shown three times in random order. After the exposure to individual patterns, participants had to complete a two-alternative forced choice (2AFC) task involving the 30 patterns from the familiarisation block contrasted against 30 foils, which were entirely novel patterns based on new CVC words (see Table 3 in Appendix A for all the foils). The co-occurrence of foil and target pattern was randomised. Participants were required to use the mouse to click on the pattern they thought they had viewed during familiarisation. Importantly, for each decision they made, they received immediate feedback. They were shown the number of correct responses after each choice, which increased with every correct decision made. The 1-block condition ended after one cycle of familiarisation and one 2AFC test on all 30 items presented during familiarisation. Participants in the 4-block condition went through four familiarisation blocks and four 2AFC tasks, and participants in the 8+2-block condition through a total of ten familiarisation blocks, each followed by a 2AFC test, before moving on to the grammar learning task.

Before starting the experiment, participants were informed that they were going to be trained and tested on individual black and white matrix patterns, and that it was important for them to learn each pattern in preparation for a subsequent experiment involving sequences of these patterns. Moreover, they were informed that the training involved feedback, which was aimed at assisting them in learning the patterns. Participants were told that, after this training on individual patterns, they would move on to a subsequent experiment involving sequences of patterns.

The grammar learning task was identical to Experiment 3 with regards to the training phase, method, materials and participant instructions. There was however, one difference: Rather than participants making grammaticality judgements on 144 three-element sequences during the testing phase, these were randomly reduced to 72 sequences for 8+2-block condition of the present experiment in order to reduce overall duration of the experiment on Day 2. As in Experiment 3, the present experiment not only involved two separate Languages L1 and L2 to control for arbitrary preferences people may have for certain patterns, but also subgroups a and b, which differed with regard to the grammatical violations participants encountered on test (see Table 2.2, p.63).

6.1.2 Results

6.1.2.1 Familiarisation phase

An illustration of mean percentage correct responses can be found in Table 6.1.

Table 6.1: Discrimination of between target and foil patterns on 2AFC test. Percentage correct responses on the final block of each condition.

Condition	% correct	SD
1-block	65.78	5.41
4-block	81.46	10.40
8+2-block	90.63	9.68

In order to gain an insight into what effect increased familiarisation has on participants' ability to discriminate between target pattern and foil, a One-way ANOVA was conducted on the percentage of correct responses on the final block of each condition (i.e. responses on the first block of the 1-block condition, on the fourth block of the 4-block condition and on the tenth block on the 8+2-block condition). Condition (1-block, 4-block and 8+2-block) was employed as a between-subjects factor. The analysis resulted a significant effect, $F(2,46) = 31.15$, $p < .001$. A subsequent Tukey post-hoc analysis showed that percentage of correct responses differed significantly between all three conditions as can be seen in Table 6.2. This therefore accounts for the significant effect.

Table 6.2: Results of the post-hoc analysis showing that performance on the final 2AFC test in each condition differed significantly.

	1-block	**4-block**	**8+2-block**
1-block	-	<.001	<.001
4-block	<.001	-	.014
8+2-block	<.001	.014	-

The ANOVA along with the post-hoc analysis therefore clearly demonstrate that the more training participants received, the more familiar they became with individual matrix patterns, as shown by the increased ability to successfully discriminate target patterns from foils. Even just after the minimal amount of familiarisation participants successfully distinguished between target and foil (see Figure 6.2). This was shown in an additional one-sample t-test, which was carried out to compare performance in the 1-block condition to 50% chance performance and revealed a significant difference ($t(14) = 11.29$, $p < .001$.

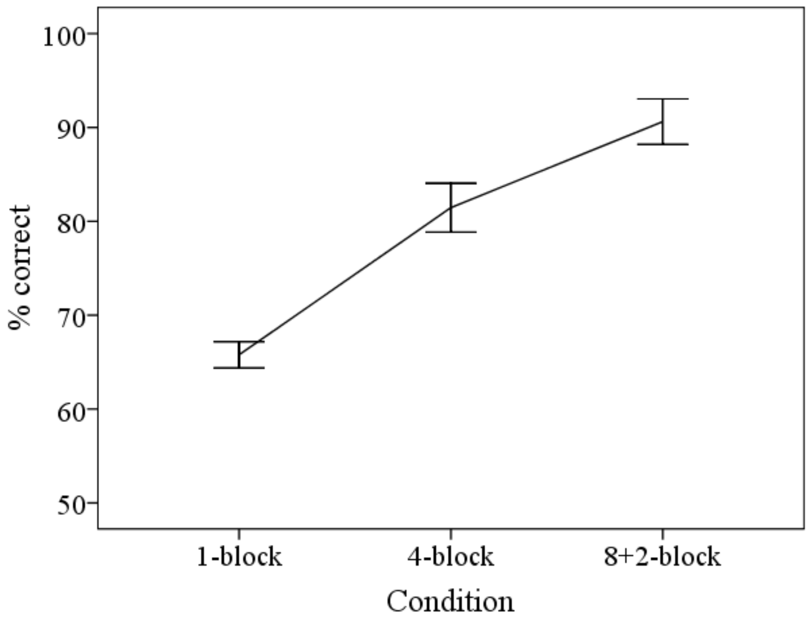


Figure 6.2: Familiarity with Componential patterns increases with the more training blocks participants receive.

6.1.2.2 Grammar learning phase

Initially, three separate one-way ANOVAs were carried out to contrast percentage correct responses between the subgroups L1a, L1b, L2a and L2b in the three conditions. The Levene's test for the 1-block condition was significant ($p = .044$), and therefore the Welch F-ratio is reported. There was no significant effect of Subgroup on the percentage correct responses, $F(3,6.38) = 4.19$, $p = .06$. The ANOVA for the 4-block condition also revealed no significant difference between subgroups, $F(3,15) = .087$, $p = .09$, as did the ANOVA for the 8+2-block condition, $F(3,15) = .889$, $p = .48$. All subgroups were therefore collapsed for further analyses. The full ANOVA results can be found in Appendix B (Table 13). Table 6.3 shows mean percentage endorsements for grammatical and ungrammatical strings containing familiar and unfamiliar X patterns.

Table 6.3: Mean percentage endorsements for grammatical and ungrammatical strings, containing familiar and unfamiliar X items, across all three conditions.

Condition	Grammaticality	Familiarity	% endorsements	SD
1-block	grammatical	familiar	60.42	12.23
		novel	52.43	13.79
	ungrammatical	familiar	57.63	7.89
		novel	49.31	12.36
4-block	grammatical	familiar	58.68	14.30
		novel	49.31	13.05
	ungrammatical	familiar	53.30	16.98
		novel	50.35	16.57
8+2-block	grammatical	familiar	65.97	20.17
		novel	43.75	21.26
	ungrammatical	familiar	60.76	24.47
		novel	44.44	25.26

As a key aspect of the present experiment was to investigate the impact that prior training on the vocabulary of the patterns has on performance in the grammar learning task, data collected for the Componential condition in Experiment 3 were included in the present ANOVA.

Importantly, the Componential condition in Experiment 3 differed from Experiment 6 with regard to the familiarisation prior to the AGL task (i.e. there was no familiarisation for participants in Experiment 3). For the present analysis, this data was included as the 0-block condition. An ANOVA was conducted on the percentage of endorsements, with Familiarity (old versus new X item) and Grammaticality (grammatical versus ungrammatical) as within-subjects factors and

Training (0-block versus 1-block versus 4-block versus 8+2-block) as a between-subjects factor[16]. The full ANOVA table can be found in Table 15 in Appendix B.

The analysis resulted in one main effect and one interaction. The significant effect for Familiarity (familiar X versus novel X), $F(1,76) = 30.52$, $p < .001$, showed that overall participants were more likely to endorse strings containing a familiar X item (57.85%) as opposed to strings containing novel X units (47.01%). Additionally, there was a significant Familiarity x Grammaticality interaction, $F(1,76) = 4.302$, $p = .041$ (see Figure 6.3). There were no further significant effects or interactions.

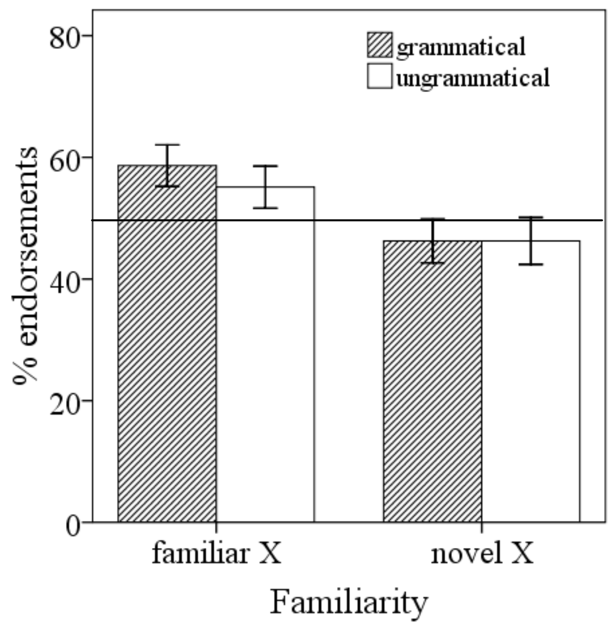


Figure 6.3: Familiarity x Grammaticality interaction, with error bars (+/- 1SE). The black line across indicates chance performance at 50%.

The interaction was further explored in two paired-samples t-tests, contrasting percentage endorsements for grammatical against percentage

[16] Note that the conditions also differed with regard to how many test items participants were exposed to during the testing phase. It can therefore not be ruled out that this may have impacted on the results.

endorsements for ungrammatical sequences, for strings containing familiar and novel X patterns separately. A corrected alpha-level (α = .025) was employed.

Table 6.4: Mean percentage endorsements for grammatical and ungrammatical strings, along with results from the paired samples t-tests, for sequences containing familiar and unfamiliar X items.

X item	Grammaticality	% endorsements	Statistic
familiar	grammatical	58.65 (15.26 SD)	t(79) = 1.75, p = .084
	ungrammatical	55.10 (15.48 SD)	
novel	grammatical	46.25 (16.11 SD)	t(79) = .00, p = 1.0
	ungrammatical	46.25 (17.25 SD)	

The t-tests suggest that discrimination between sequences based on grammaticality was unsuccessful for strings containing familiar as well as for strings containing novel X items. Thus, in line with previous analyses, the difference scores (i.e. percentage endorsements with grammatical – percentage endorsements with ungrammatical sequences) were calculated along both levels of familiarity with X, and submitted to a paired-samples t-test. This t-test showed that discrimination between sequences based on grammaticality was significantly better for sequences involving familiar X patterns (3.54%, 18.11 SD) as opposed to novel X items (.00%, 17.57 SD), t(79) = 2.08, p = .040. This difference therefore accounts for the interaction. However, since grammaticality judgements were unsuccessful for sequences with novel as well as with familiar surface forms, this interaction does not reflect a meaningful effect.

Of particular importance is also the absence of a main effect for Grammaticality, F(1,76) = 1.12, p = .294, as well as the absence of a significant Training x Grammaticality interaction, F(3,76) = .08, p = .972, showing that performance based on the grammaticality of the sequences did not differ depending on the amount of training participants received. Thus, the underlying grammar was not identified in any of the conditions in this experiment.

6.1.3 Discussion

The aim of Experiment 6 was two-fold. Its first aim was to familiarise participants with the complex black and white matrix patterns used for the Componential conditions in experiments 3 and 4. The second goal of this experiment was to manipulate the amount of familiarisation the participants received in order to investigate whether this would impact on their ability to detect the non-adjacent dependencies during the subsequent grammar learning experiment.

The present findings indicate that the more familiarisation blocks participants received, the more likely they were to successfully discriminate between familiar and entirely novel patterns on test. There was thus a definite learning effect, and increased training duration lead to increased familiarity with Componential patterns. Furthermore, increased familiarisation also reduced confusability between complex patterns as people's performance on the 2AFC test increased from approximately 65% correct responses (1-block) to approximately 80% (4-block) to around 90% in the 8+2-block condition.

However, people's competence at discriminating between target pattern and foil did not affect their performance during the AGL task. Here, participants failed to extract the non-adjacent dependencies and base their grammaticality judgements on the underlying PSG. The absence of a significant Training x Grammaticality interaction showed that this was true even for participants in the 8+2-block condition, who were significantly more familiar with the patterns than people in the other two conditions.

In general then, this experiment provides evidence that the individual Componential patterns are learnable and distinguishable from one another. Yet participants are unable to learn the non-adjacent dependencies even when they are highly familiar with individual patterns. A potential explanation for the present findings might be that simultaneously presented sequences of three matrix patterns lead to a computational explosion, and participants were overwhelmed by the sheer number of potential part-to-part and whole-to-whole relationships available during each trial in the AGL task of the experiment. However, the main effect for Familiarity made clear that participants favoured familiar patterns, and thus must have recognised familiar patterns, even when they were exposed to sequences of three. Thus, they were capable of analysing the patterns on an individual basis and in sequences of three, yet they were unable to extract the underlying grammar generating the sequences.

With regard to the potential reasons why participants fail at learning non-adjacent regularities between componential patterns as identified in the introduction of this chapter, the possibility of familiarity with the matrix patterns and confusability between matrix patterns can be, to an extent, excluded based on the present experiment. At the same time, however, all the participants in this experiment have had much more exposure to language than to the black and white matrix patterns. There is therefore still a possibility that further increasing the prior familiarisation to the patterns before the AGL task would result in the non-adjacencies being learnt.

The fact that the main effect for Familiarity revealed that participants base their grammaticality judgments on the surface form of the sequences during the AGL task is interesting. This may indicate that extensive exposure to Componential patterns leads to stimulus-specific knowledge. The emergence of knowledge of the specific experimental stimuli as opposed to knowledge of abstract representations underlying the stimuli has been found in previous work involving computational models (Christiansen & Curtin, 1999; Dienes, Altmann & Gao, 1999). More recently, Johansson (2009) carried out three experiments to investigate this effect in an AGL paradigm. Of particular relevance in this context is Johansson's third experiment, in which he used only visual stimuli, specifically simultaneously presented sequences of abstract shapes or of different colours, generated by a FSG. Crucially, he manipulated the amount of training participants received so that one condition involved a short training phase and the other condition involved a long training phase. The results showed that participants with longer training phases and thus longer exposure to the visual stimuli based their grammaticality judgments more on the stimuli themselves than participants with shorter training phases. In a similar fashion, the preceding familiarisation with the Componential vocabulary in the present experiment resulted in good stimulus-specific knowledge in that the patterns themselves were learnt (or at least recognised). Even in the 0-block condition, the AGL paradigm involved a fairly extensive training phase, and sequences containing patterns were clearly favoured over sequences with novel surface structures (i.e. with novel X items). Importantly, this extensive exposure did not enable participants to abstract the underlying grammar of the sequences, i.e. learn the grammar and apply it to novel instances. The present experiment therefore confirms that increased exposure results in stimulus-specific knowledge.

Thus, taken together, the present findings can be seen as supporting evidence for the language-specific knowledge posited in the previous two chapters, as non-adjacent dependencies between Componential patterns are not learnable even after extensive prior familiarisation. The reason why the non-adjacent regularities are not acquired between Componential patterns may therefore be due to the inability to analyse two levels of encoding in the non-linguistic domain. In the linguistic domain, however, this duality of patterning is more readily acquired. The language-specific knowledge, i.e. the knowledge of which regularities matter in language, which has been identified step by step throughout the work reported in this thesis may well represent a subsystem within a larger, domain-general system that subserves the acquisition of non-adjacent dependencies.

The notion that certain cognitive processes, specifically within the area of sequence learning, are not guided by just one general-purpose mechanism is not entirely new. Conway and Christiansen (2006) argued for modality-specific representations on the basis that they found participants to be able to learn two separate grammars simultaneously only if they were instantiated by two different perceptual modalities. When the two grammars were represented by perceptually similar materials, then a learning effect was found only for one of the grammars. They conclude that statistical learning is closely associated with the modality of the input, and that separate modality-specific subsystems are capable of operating independently of each other (and thus learning two grammars across two modalities), yet one of these subsystems is incapable of extracting two grammars simultaneously. This research along with the present findings highlights the possibility that different sequence learning abilities involve task-specific, stimulus-specific, modality-specific or domain-specific subsystems. These individual subsystems that lie at the core of learner's cognitive processes have also been suggested by Conway and Pisoni (2008), Conway and Christiansen (2005) and Goschke (1998). This issue will be discussed in further detail and related to the larger research area in the following chapter.

6.2 Concluding remarks

A very refined combination of global and local analysis is required in order to detect non-adjacencies between componential patterns: The sub-patterns matter to make the patterns distinguishable, yet the patterns matter for the acquisition of the non-adjacent grammar. The present results show that people are capable of the local analysis, as there was a definite learning effect in the Familiarisation task, yet they still fail to detect the global regularities. Thus, when the internal structure of the non-linguistic materials are designed to replicate a defining feature of language, as here with the Componential materials, then the non-adjacencies are not identified. The present findings were therefore interpreted as potential evidence for the language-specific knowledge put forward in Chapter4 and 5. The overall findings of the entire series of experiments presented in this thesis are discussed in detail and related to other relevant literature in the following chapter.

Chapter 7

General discussion

Natural languages are structurally immensely complex and yet children acquire their native language with relative ease (Chomsky, 1965). Bridging this gap requires either the assumption of an innate LAD (Chomsky, 1965), which is designed specifically for language acquisition, or the existence of efficient domain-general learning tools that can be recruited for language learning. Recent AGL and ALL paradigms have identified a number of mechanisms employed for language learning (Section 1.3), and have explored to what extent these operate across domains (Sections 1.3.1 – 1.3.4).

7.1. Summary

The aim of the present thesis was to investigate the acquisition of non-adjacent dependencies. Non-adjacent dependencies represent a structural complexity frequently found in natural languages, and crucially, they are a diagnostic of PSGs. As such they form an excellent test case for investigations into the domain-specificity of the human linguistic capacities.

With regard to the identification of non-adjacent dependencies in the linguistic domain, the work carried out for the present thesis does not confirm previous research, which suggest that non-adjacencies are unlearnable unless additional cues are provided. Specifically, Experiment 1 and Experiment 2 investigated two types of cues. Previously, the Gestalt Principle of Similarity has either explicitly been proposed as a requirement for the successful detection of non-adjacent regularities (Newport & Aslin, 2004) or has been inadvertently included in the stimulus materials (Gómez, 2002; Peña, Bonatti, Nespor & Mehler, 2002). The results from Experiment 1 do not corroborate these findings as it was shown that the Gestalt cue did not significantly enhance the extraction of the non-adjacent dependencies. As the Gestalt cue can be interpreted as a way of highlighting the relevant elements of analysis, Experiment 2 addressed a further cue, which could also have highlighted the crucial artificial words. Here, break cues were manipulated, and the results showed that neither the elimination of helpful inter-word breaks, nor the inclusion of unhelpful inter-grapheme breaks rendered non-adjacencies unlearnable. Thus, the acquisition of non-adjacent dependencies is not dependent on additional, domain-general cues (as in the Gestalt principle of Similarity) or non-linguistic cues (as in the breaks).

As the main focus of the thesis regards the extent to which the detection of non-adjacent dependencies is possible across domains, the remaining experiments focused on the contrasting performances in the linguistic and the non-linguistic domain. In general, the findings support the idea of language-specific expectations, meaning that people are guided by their prior knowledge of language when processing a linguistic input. Experiments 3 and 4 found that the non-adjacent dependencies in the non-linguistic domain were only learnable when the correct level of analysis was highlighted. This was achieved by using simple, unanalysed shapes as stimulus materials. Crucially, when the non-linguistic materials contained internal structure, and specifically exhibited systematic re-use of sub-units in a way that language does, then the non-adjacent dependencies were not detected. This was the case regardless of whether the sequences were presented simultaneously (Experiment 3) or sequentially (Experiment 4).

Experiment 5 investigated the domain-specificity of the acquisition of non-adjacent dependencies in the auditory modality, and found that like in the visual modality, the non-adjacent dependencies between words were reliably detected in the linguistic domain. Interestingly, non-adjacencies were most reliably extracted in the –Gestalt condition, confirming that the Gestalt Principle of Similarity does not facilitate the detection of non-adjacencies. In the non-linguistic domain, the materials were instantiated as sounds differing in timbre as it was argued that the stimuli used in previous experiments (Creel, Newport & Aslin, 2004) may have interfered with people's expectations regarding musical tones. This, however, resulted in the sounds being less informative than in the visual modality, and this is presumably the reason for participants' relative poor performance in the auditory equivalent of the Shape condition. Importantly, the regularities were not detected in the auditory Componential condition.

The final experiment ruled out the possibility of participants' unfamiliarity with the Componential patterns impacting on their ability to detect non-adjacent dependencies, as the learning effect found in the Familiarisation phase did not carry over into the AGL task. This is therefore interpreted as a confirmation for the existence of language-specific knowledge. It is argued that this language specific knowledge guides people toward the regularities that matter in language, and the present experiments seem to confirm that this knowledge does not operate (or is not used) in the non-linguistic domain. In this way, the work presented in this thesis is counter to the general consensus which argues for mainly domain-general mechanisms underlying the human capacity for language (Perfors, Tenenbaum, Regier, 2010).

The results suggested a need for a more refined view of the systems subserving the human language capacity. The results were therefore discussed in accordance with Atkinson and Wheeler (2004). It was argued that the mechanism recruited for the detection of non-adjacencies might be domain-general, as the non-adjacencies were acquired in the Shape, +Gestalt and –Gestalt conditions. At the same time, this mechanism must be capable of drawing from language-specific knowledge, as the regularities were not extracted in the Componential conditions, which replicated the internal structure of language. These language-specific expectations, established throughout this thesis, therefore constitute the main theoretical contribution to the research area. It is argued that when exposed to linguistic stimuli, people's knowledge of what regularities matter in language will guide their attention toward the relevant structural elements or level of analysis. The non-adjacencies in all linguistic conditions were reliably detected, as relationships between elements separated by intervening units are frequently found in English. When the internal complexity of language is mimicked in the non-linguistic domain, however, this knowledge is either not accessible or not used to analyse the input, which results in complex structural regularities, such as non-adjacent dependencies, not being detected. This was demonstrated in the Componential conditions, where participants were unaware of the regularities holding across the first and third patterns or sounds. However, the detection of non-adjacent dependencies is not unique to language, as participants successfully acquired the grammar in the visual Shape condition. Thus, when the internal complexity is removed from the non-linguistic materials, and the relevant units of analysis are thereby highlighted, then people are capable of detecting non-adjacent dependencies in the non-linguistic domain.

However, the present experiments cannot provide an entirely clear account of the correct level of explanation. To make the argument clearer, let’s assume that the mechanism underlying the acquisition of non-adjacent dependencies is domain-general (ignoring the language-specific expectations for a moment). This could mean either that one and the same mechanism operates across both the linguistic and non-linguistic domain, or that one mechanism, say a language-specific mechanism, has been replicated and operates separately in the non-linguistic domain. Crucially, the present results can give no insights into which is the case. This will therefore be addressed in 7.3.

7.2 Methodological concerns

From a methodological perspective, there were difficulties involved with conducting the present series of experiments. The first difficulty regards the adequate translation of stimulus materials from one modality to the other. As language can be easily instantiated as written text in the visual modality and as speech (or in this case as synthesised speech) in the auditory modality, translating the +Gestalt and –Gestalt stimuli from the visual to the auditory modality was fairly easy. For the non-linguistic domain, however, these cross-modal investigations were much more difficult. Finding equivalent auditory stimuli for the visual Shape condition was particularly challenging. In the visual modality, the abstract shapes differed from each other along two dimensions (shape and colour), and were therefore very distinct. In the auditory modality, using sounds that differed merely in timbre may well have produced alien sounding stimuli. The sounds used for the auditory modality and the shapes used for the visual are therefore not adequate equivalents. Analysing sequences containing these sounds may thus well have resulted in much higher processing cost, and therefore inhibited the detection of non-adjacent dependencies. This highlights the difficulties involved in exploring domain-specificity across modalities using an AGL paradigm.

Secondly, contrasting people's ability to extract non-adjacent dependencies across both modalities (visual and auditory) is in itself problematic as the modalities themselves are different. In practical terms this means, for example, that participants in the visually presented experiments had the opportunity of not paying attention during the training as they could have simply chosen not to look at the monitor. In the auditory modality, by contrast, participants had no choice but to listen to the sounds.

Thirdly, the present experiments do not take into account individual differences. In normally developing adults, there are always differences with regards to their linguistic abilities, and it has been shown recently that this is also the case in AGL experiments (Misyak, Christiansen & Tomblin, 2010). By employing a novel experimental paradigm, which combines a standard AGL experiment with a serial reaction time task, Misyak et al. were able to measure online learning of non-adjacent dependencies. Specifically, their participants were exposed to aural instantiations of Gómez's (2002) AL. On hearing each word from the AL, participants were required to use the mouse to click on to the corresponding written word on the computer screen. Each word appeared on the screen with a distractor word. In Gómez's design, as in the present experiments, the non-adjacent dependencies operated between the first and the third word of each sequence. Therefore the first word reliably predicts the third word. By measuring reaction times, Misyak et al. found that in their novel experimental set-up, good learners were quicker to click on the third word of each sequence than poor learners, since they had extracted the underlying rule better. On hearing the first word of each sequence, they were aware of what the third word for the same sequences would be, and this reduced the reaction times for clicking onto the third word. This provides strong evidence for definite individual differences in AGL experiments.

Furthermore, conducting Experiment 6 revealed two specific issues. Firstly, including an additional familiarisation phase prior to the AGL task resulted in a fairly long experiment. Although the testing phase at the end of the AGL task was shortened in order to address this, the entire experiment – especially the 8+2-block condition, which stretched over two days – was fairly long. This might well have resulted in fatigue, boredom or reduced motivation on behalf of the participants, which in turn may have affected performance. The second issue regards the length of the familiarisation phases. Language is hugely overlearnt as people are exposed to it on a daily basis. People are thus highly familiar with processing a linguistic input. In Experiment 6, even the longest familiarisation phase, which involved a total of 10 familiarisation blocks, will not have familiarised participants with the complex patterns to a level equivalent with language. Therefore the processing cost involved in analysing the sequences containing three complex patterns in the AGL task would still have been much higher than for the linguistic conditions. This may, in part, explain people's inability to extract the non-adjacent dependencies.

Conducting the present series of experiments thus involved a number of methodological concerns. However, the findings nevertheless form a solid basis for future research. This will be discussed in the following section.

7.3 Future directions

There are a number of different directions for the future. As pointed out in 7.1, the present findings can only offer a limited insight into the correct level of explanation regarding domain-specificity. On the surface of it, experiments on transfer effects may seem to be an efficient way of tackling this issue. Dienes, Altmann and Gao (1999) note that transfer effects are due to participants abstracting the grammar learnt during the training phase of an AGL task, and applying this grammar to novel instantiations. Thus, if the same general-purpose learning device operates across domains, then knowledge acquired in one domain should be easily transferable to another domain. If, however, two identical mechanisms operate independently in separate domains, then this transfer should be impossible. According to this, it might seem reasonable for future work on non-adjacent dependencies to focus on exploring transfer effects from the linguistic to the non-linguistic domain and vice versa. Thus, participants would receive training in one domain, e.g. in the –Gestalt condition, and be tested on the underlying grammar in the non-linguistic domain, e.g. in the Shape condition. However, there is a difficulty with this design: Transfer effects have been shown to be carried by repetition structure (Brooks & Vokey, 1991; Gómez & Gerken, 2000; Mathews & Roussel, 1997, Tunney & Altmann, 1999). Crucially, the PSG employed in the present series of experiments does not involve repetition structure as none of the elements was repeated within any of the sequences. The reason for this was that the present experiments explored the acquisition of non-adjacent dependencies between categories rather non-adjacent repetition of identical elements. Thus, investigating the acquisition of non-adjacencies on the syntax level, governed by a PSG, does not lend itself easily to transfer experiments as it does not include repetitions of specific elements.

A better way of taking the present results one step further and to explore the refined view of domain-specificity argued for here, are AGL experiments involving Event Related Brain Potentials (ERPs). Conducting ERP experiments in order to pinpoint functional areas underpinning language is difficult due to the sheer complexity of language, especially syntax (Pulvermüller, 2010). For this reason, isolated linguistic computations, for example specific syntactic operations, have to be clearly identified prior to the ERP investigations. There are two typical brain signatures elicited during syntactic processing in natural languages. Firstly, there is a late positivity, the P600 effect, which is elicited when processing a range of syntactic anomalies, such as agreement violations (as in **The socks is*) or category violations (as in **I like of*) (Friederici, Pfeifer & Hahne, 1993; Hagoort, Brown & Groothusen, 1993; Lau, Stroud, Plesch & Phillips, 2006). Secondly, the ELAN effect, which is an early left anterior negativity, is associated with a more restricted range of syntactic violations, and is thus elicited mainly due to category violations (Hahne & Friederici, 1999; Lau, Stroud, Plesch & Phillips, 2006). In an AGL experiment involving ERPs, Freiderici, Steinhauer and Pfeifer (2002) found that their participants exhibited similar patterns of brain activation as would be expected in natural language processing. Participants in their design, in contrast to the majority of AGL experiments, received extensive training on the AG over a number of sessions, with one session lasting up to five hours. Additionally, the training did not involve mere exposure to the grammar as in most other AGL experiments (including virtually all the experiments discussed in this thesis), but involved participants communicating and interacting with each other using the AL. Participants were tested on grammatical and ungrammatical sentences, the latter of which contained category violations only. Their results demonstrated that participants trained on the AG showed an ELAN signature as well as a P600 effect, as would have been expected in natural language processing. Similar results were found by Christiansen, Conway and Onnis (2007). This demonstrates that AGL paradigms can be combined with ERP experiments to investigate brain activation. In a related design, Baldwin and Kutas (1997) found that when trained and tested on a light moving around a 9-

square grid in a predetermined fashion (governed by a FSG), their participants did not show an ELAN or a P600 effect. This may indicate that the neural circuits recruited for grammar learning in the linguistic domain are different to the ones recruited for the non-linguistic domain. However, since Friederici et al. and Baldwin and Kutas's experiments differ across a number of dimensions, most notably in the structure of the AG, no direct comparisons can be drawn. For this reason, an interesting future direction for this line of research is to adapt the present AGL experiments for an ERP experiment. This could give an indication as to whether the same neural mechanism is employed for the detection of non-adjacent across the linguistic and non-linguistic domain, or whether there are in fact two independent mechanisms, one for the linguistic and one for the non-linguistic domain.

Moreover, another potentially fruitful direction to take with this line of research is to study the acquisition of non-adjacencies in children and infants. All the experiments presented here were conducted on adult participants as the main focus here was domain-specificity. The findings for the linguistic conditions have shown that non-adjacencies between words are reliably detected, even without assistance of Gestalt cues. For this reason, it would be interesting to carry out a slightly adapted version of the present AGL experiments on infants. For the linguistic domain, the experiments would have to be conducted in the auditory modality. Gómez (2002) tested 18-month old infants on two dependencies (as opposed to three), and up to a set size of 24 X items. This age group would therefore be the ideal starting point to investigate whether the Gestalt cues, as included by Gómez, is indeed necessary for the detection of non-adjacencies in the linguistic domain. The AL would have to be reduced accordingly. The next step would then be to explore domain-specificity. For the Shape condition, the visual modality would be preferable as the sounds used for the present auditory modality rendered unsatisfactory results, and because Kirkham, Slemmer and Johnson (2002) showed that infants as young as 2 months can successfully take part in a visual sequence learning task, by monitoring looking times. These experiments would give an insight into whether and to what extent non-adjacencies are readily extracted during language acquisition. Based on Gómez's (2002) findings as well as the present results, the prediction would be that 18-month old infants would be able to extract non-adjacent dependencies in the –Gestalt condition, using 24 intervening X items. Particularly in the present auditory –Gestalt condition, adult participants reliably detected the regularities, and therefore a similar finding would be expected with 18-month olds. However, due to the fact that tracking non-adjacent dependencies is cognitively more demanding (Santelmann & Jusczyk, 1998), younger infants may well struggle. For the Shape materials, the prediction would not be as clear. Based on performance levels in the present visual Shape conditions, it might be expected that 18-month olds would be able to extract the regularities. However, Johnson, Fernandes, Frank, Kirkham, Marcus, Rabagliati and Slemmer (2009) found that 8- and 11-month old infants who were trained on an ABA pattern failed to detect the non-adjacent repetitions on test. This may be an indication of the cognitive strain that non-adjacencies pose in conjunction with the fact that Johnson et al. used younger infants than Gómez did. Another (speculative) explanation may be that this reflects the fact the detection of non-adjacencies is primarily possible in the linguistic domain, and only possible in the non-linguistic domain later on in development. Thus, conducting the present AGL experiments on infant participants would potentially produce highly interesting and insightful results.

The novel hypothesis established throughout this thesis is that domain-specific expectations feed into a domain-general learning device, meaning that people have language-specific expectations about the relevant units of analysis in the linguistic domain. An interesting way of testing this hypothesis would be to conduct cross-linguistic experiments using populations that may be equipped with different linguistic expectations than native speakers of English. The point was made earlier (Experiment 3) that the re-use of segments in the linguistic conditions did not distract people away from the non-adjacent dependencies as regularities between segments is not of importance in the English language. The systematic re-use of sub-patterns in the Componential conditions, so it was argued, did distract people and the non-adjacencies were therefore not learnable. However, some languages, such as Finnish, display Vowel Harmony, meaning that regularities between segments here do play an important role. It would therefore be interesting to run the –Gestalt condition on an experimental group of native speakers of Finnish and a control group of English speaking participants. If it is indeed a case of language-specific expectations guiding people toward the correct analysis, then the Finnish-speaking participants would be expected to show some kind of interference when it comes to extracting the non-adjacent dependencies based on the fact that they are used to Vowel Harmony. Vowel sets in Finnish determine grammaticality. If the initial vowel is a back vowel, all subsequent vowels must be either back or neutral. If the initial vowel in a word is a front vowel, all subsequent vowels must be either front or neutral. The back vowel set includes /u, o, a/, front vowels are /y, ø, æ/ and neutral vowels are /i, e/ (Suomi, McQueen & Cutler, 1997). The experiment would have to be conducted in the auditory modality as some graphemes used in Finnish are not used in English. Also, the present AL would have to be somewhat modified. For example, in L1a of the –Gestalt materials, the grammatical non-adjacent dependencies are *lum* X *fip*, *zel* X *pof* and *vok* X *gam*, and the ungrammatical test items involve **lum* X *gam*, **zel* X *fip*, and **vok* X *pof*. This would have to be changed to the grammatical dependencies being *lum* X *fyp*, *zæl* X *pof* and *vok* X *gam*, and the ungrammatical strings being **lum* X *gam*, **zæl* X *fyp*, and **vok* X *pof*. Crucially, all words in category X would have to be CVC words containing only the neutral vowels /i/ or /e/. Importantly, in all three ungrammatical sequences for this experiment, the non-adjacent items would contain vowels that are taken from the same vowel set. If Finnish-speaking participants were guided by their prior linguistic knowledge, then it might be expected that this knowledge would interfere with their grammaticality judgments, specifically with their ability to reject sequences that conform with Vowel Harmony and accept sequences that do not. If this were indeed the case, then these findings could be interpreted as supporting evidence for language-specific knowledge.

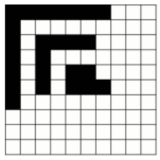
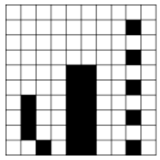
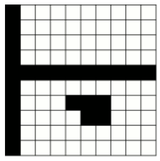
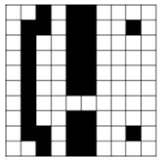
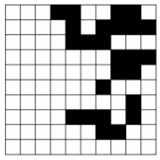
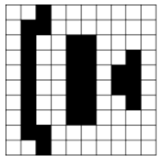
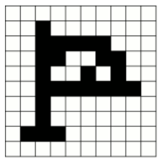
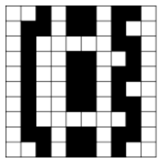
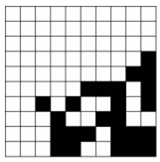
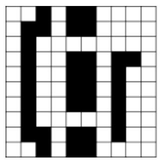
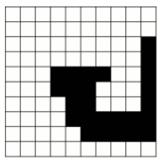
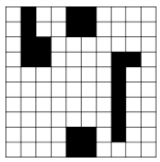
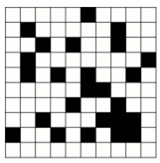
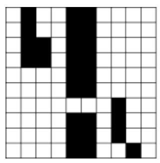
An additional interesting way of testing this language-specific knowledge would be to run the present AGL experiment using linguistic characters, which are unfamiliar to UK students, such as Cyrillic symbols. There would have to be two separate experimental conditions: One in which participants are explicitly told that they were going to see sequences of language (albeit a foreign language using unusual characters), and one condition in which participants are merely told they are about to see a large number of sequences made up of three abstract symbols. It would be interesting to see whether people's explicit knowledge of whether or not they are processing linguistic sequences affects their ability to detect the non-adjacent dependencies. In line with the hypothesis established in this thesis, the prediction would be that participants who are not informed about the domain of the input would struggle to detect the underlying regularities, and be outperformed by participants in the other condition.

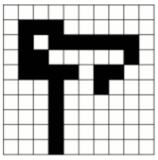
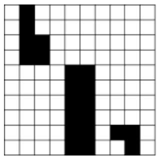
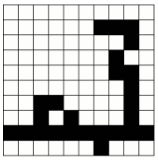
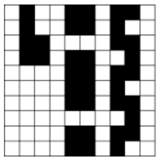
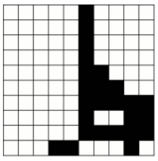
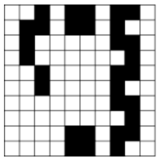
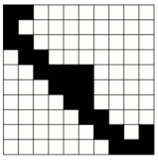
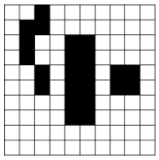
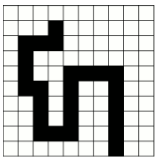
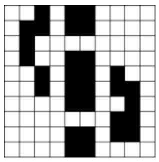
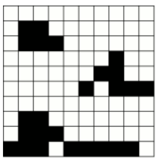
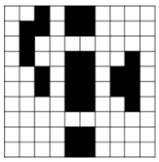
To summarise, the work carried out for this thesis expands on the present knowledge of learning tools available to humans in language learning. Specifically, it found that the acquisition of non-adjacent dependencies in the linguistic domain is robust and not reliant on additional cues. This is in line with what was expected initially as non-adjacent regularities are not specifically highlighted in natural languages. The present series of work showed that the detection of non-adjacencies is a domain-general ability, which, to an extent, is specialised for linguistic analysis. Moreover, these experiments form a good basis for a number of interesting follow-up investigations.

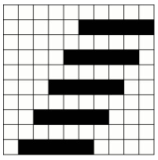
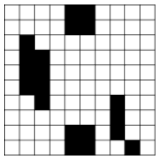
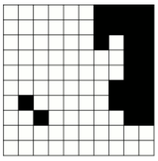
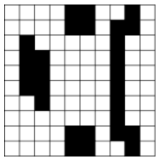
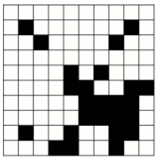
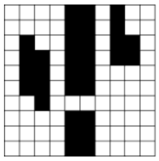
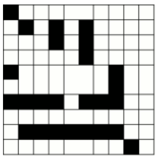
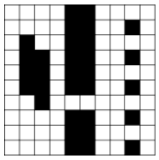
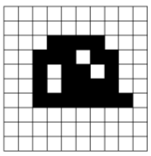
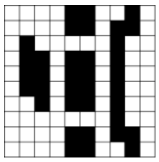
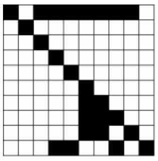
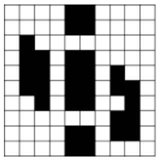
Appendix A

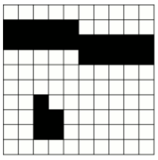
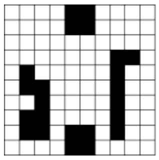
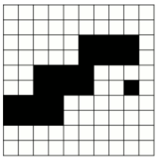
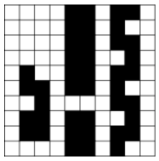
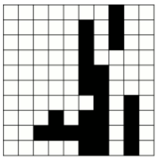
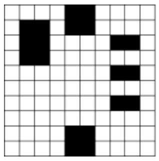
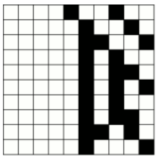
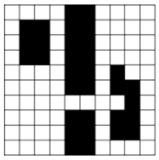
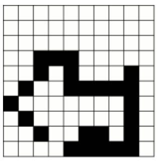
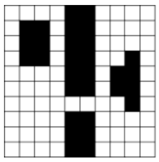
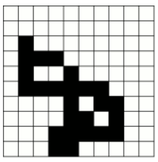
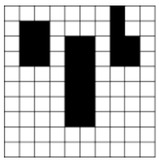
- Stimulus materials used for Experiments 3, 4 and 6: Table 1
- Logic® soundnames used for Experiment 5: Tables 2A and 2B
- Phonetic transcriptions for artificial words used in Experiment 5: Table 2C
- Foils used 2AFC task in Experiment 6: Table 3

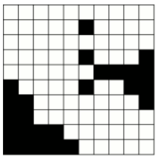
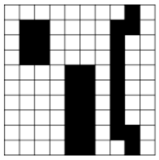
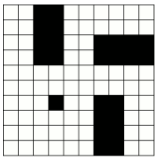
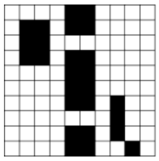
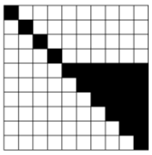
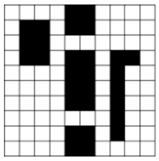
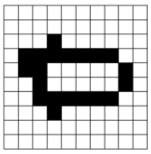
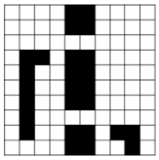
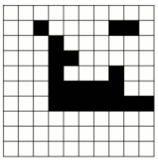
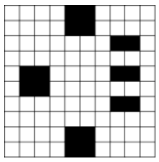
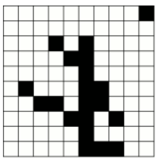
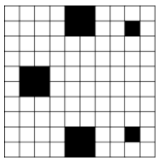
Table 1: Visual stimuli used for Experiment 3 (Shape, Holistic, Componential), Experiment 4 (Shape, Componential) and Experiment 6 (Componential).

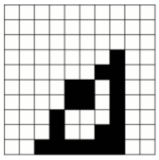
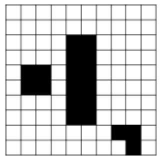
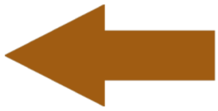
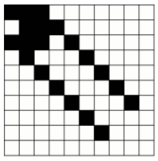
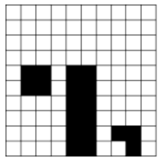
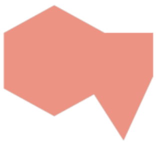
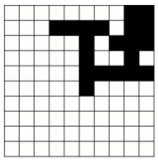
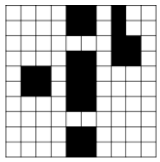
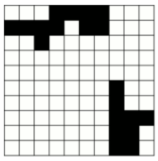
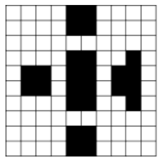
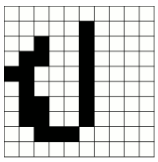
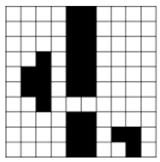
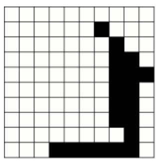
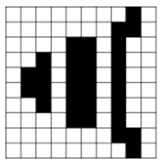
-Gestalt	Shape	Holistic	Componential
dov			
fet			
fip			
fub			
fum			
gam			
ged			

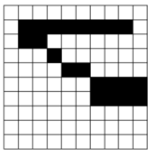
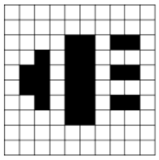
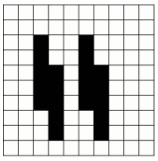
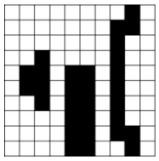
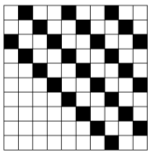
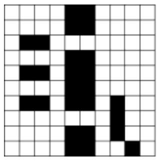
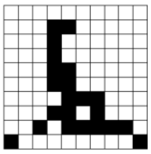
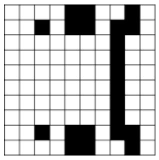
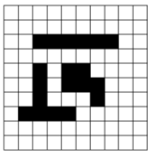
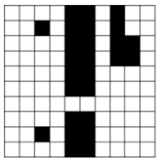
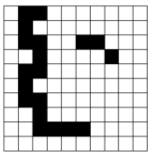
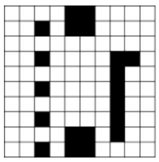
-Gestalt	**Shape**	**Holistic**	**Componential**
gos			
gub			
hab			
hin			
huk			
hup			

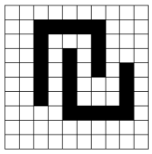
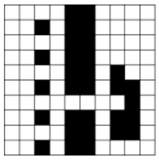
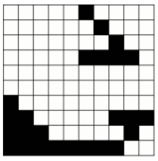
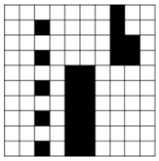
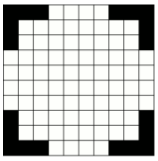
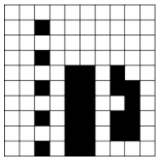
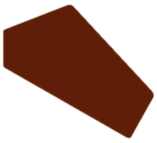
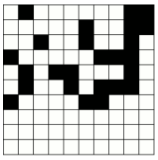
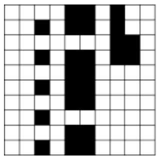
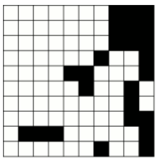
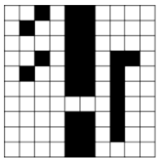
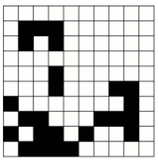
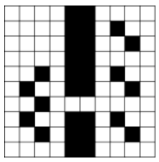
-Gestalt	Shape	Holistic	Componential
jad			
jaf			
jeg			
jev			
juf			
juk			

-Gestalt	Shape	Holistic	Componential
kam			
keb			
lar			
lek			
lep			
lig			

-Gestalt	Shape	Holistic	Componential
lof			
lud			
lum			
mus			
nar			
nat			

-Gestalt	**Shape**	**Holistic**	**Componential**
nis			
nos			
nug			
nup			
pes			
pif			

-Gestalt	Shape	Holistic	Componential
pir			
pof			
rud			
taf			
teg			
vam			

-Gestalt	Shape	Holistic	Componential
vek			
vog			
vok			
vug			
wem			
zec			

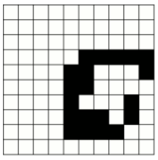
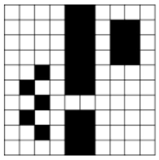
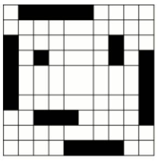
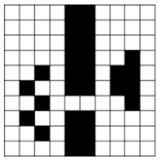
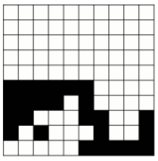
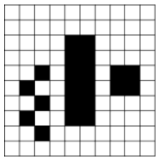
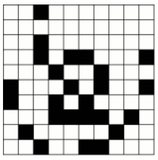
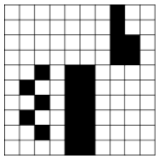
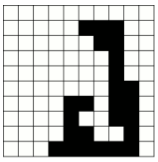
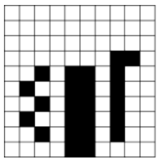
-Gestalt	Shape	Holistic	Componential
zel			
zep			
zin			
zog			
zom			

Table 2A: Grapheme-soundname mappings for the Componential condition in Experiment 5.

Grapheme	Logic Soundname
a	Steinway Piano
b	Full Strings Legato
c	Oboe Solo Legato
d	Hard Stage Mkll – JB Style
e	EVD Clav
f	Heavy Noise
g	Liverpool Bass
h	Full Brass Legato
i	Crunchy Funk Piano
j	House Bass
k	OVATI HARM4
l	-8VA TRUMPTS
m	Tuba Solo Legato
n	Finger Nylon
o	Love’n Organ
p	BAN DI P C
r	Tubular Bells
s	Ebony Flute
t	Upright Jazz Bass
u	70’s Funk Clav
v	Clarinet One
w	Vibraphone
z	Timpani Single Strokes

Table 2B: Word-soundname mappings for the Shape condition in Experiment 5.

Artificial Word	Logic Soundname
dov	AC Bass – True Stacc. - A
fet	CLARINET ONE
fip	EBONY FLUTE
fub	BR. DARK FADE
fum	SLO HAMMOND
gam	Marimba
ged	House Bass
gos	Tubular Bells
gub	Wurlitzer 200A Tremolo
hab	Auto-Wah Clav
hin	Clarinet Solo Legato
huk	EVB Organ
hup	Endless Cho
jad	BASS FING 16
jaf	Oboe Solo Legato
jeg	Kronky Organ
jev	VLNS 34 F
juf	Harpischord
juk	Clean Electric Guitar 1
kam	Vintage Wah Clav
keb	Piccolo Legato
lar	60’s Classic Rock Organ
lek	Upper Keyed Clav
lep	OVATION P 32
lig	EVD Clav
lof	Full Strings Legato
lud	Heavy Noise
lum	Steinway Piano
mus	Classic Solo Organ

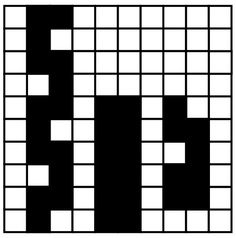
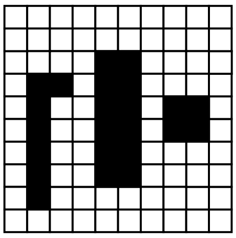
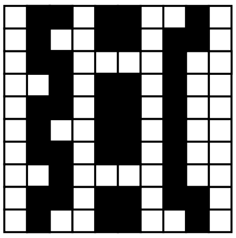
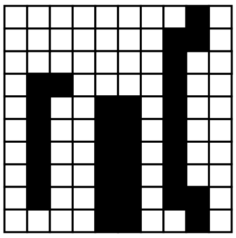
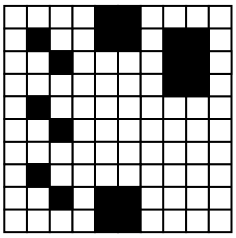
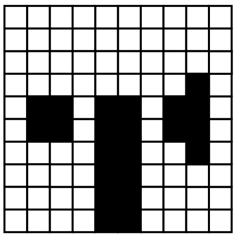
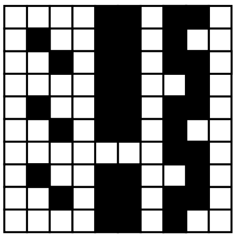
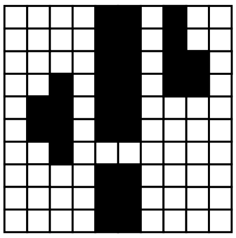
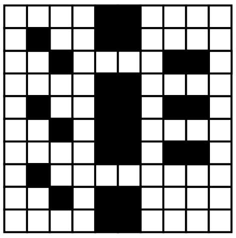
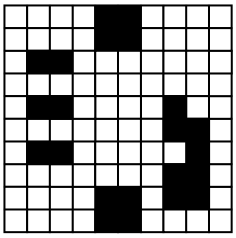
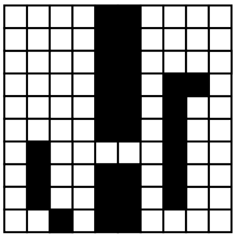
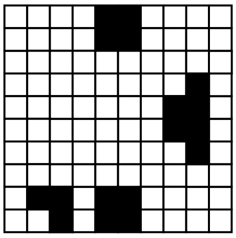
Artificial Word	Logic Soundname
nar	Classical Flute
nat	English Horn Solo
nis	Violins 1 Legato
nos	Bossa Organ
nug	70’s Funk Clav
nup	12STRING 16
pes	OVATI HARM 4
pif	Wurlitzer
pir	BAN-DI P C
pof	VIBRAPHONE
rud	Vibraphone
taf	8 Finger Nylon
teg	Timpani Single Strokes
vam	-8VA TRUMPTS
vek	CONCERT HARP
vog	HardStage Mkll – JB Style
vok	-8VA BR. SOPRA SAX
vug	Tuba Solo Legato
wem	VLAS ARCO
zec	ST JUMBO 16
zel	13BRT CELLO
zep	Cathedral - Full
zin	Full Brass Legato
zog	Liverpool Bass
zom	Love’n Organ

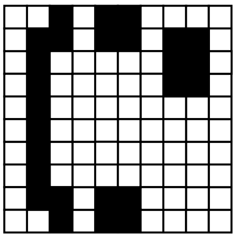
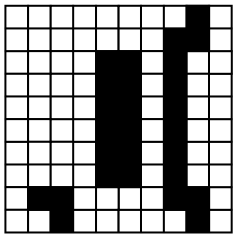
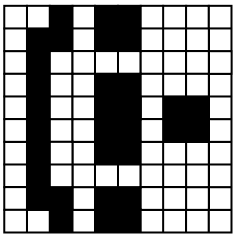
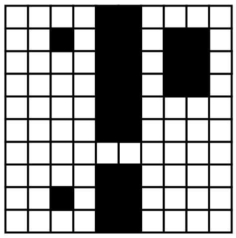
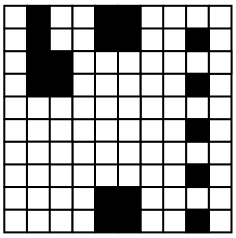
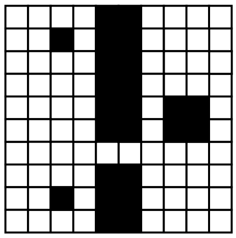
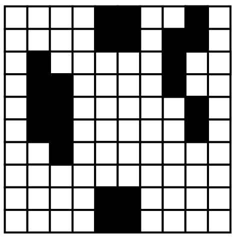
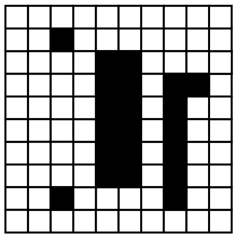
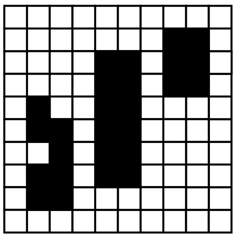
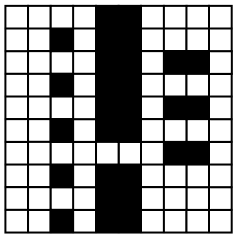
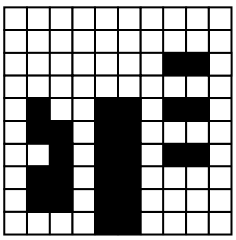
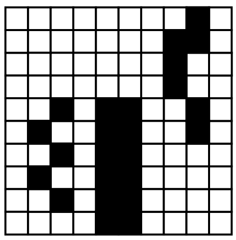
Table 2C: Phonetic transcriptions and corresponding artificial words used for the +Gestalt and –Gestalt conditions in Experiment 5. Primary stress is indicated for bisyllabic words.

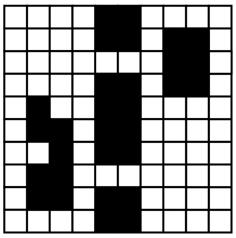
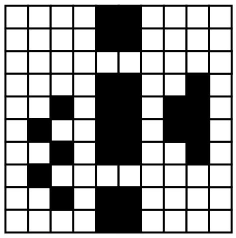
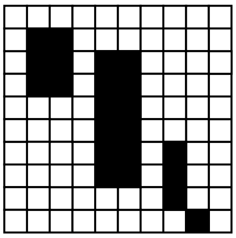
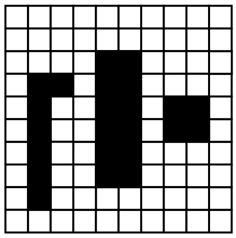
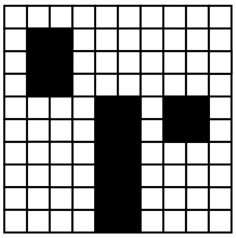
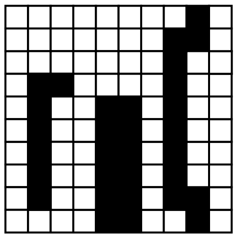
monosyllabic CVC words		**bisyllabic CVCVC words**	
[dɒv]	dov	['dəʊvæd]	dovad
[fet]	fet	['fetæk]	fetac
[fɪp]	fip	['fɪpul]	fipul
[fʌb]	fub	['fəbul]	fubal
[fʌm]	fum	['fəɒmks]	fumox
[gæm]	gam	[gæ'mʌk]	gamuc
[ged]	ged	['gɛdɒk]	gedok
[gɒs]	gos	[gəʊ'seg]	goseg
[gʌb]	gub	['gu:bɪp]	gubip
[hæb]	hab	['hæbek]	habec
[hɪn]	hin	[hɪ'nəg]	hinug
[hʌk]	huk	['hjukɪg]	hukig
[hʌp]	hup	['hu:pet]	hupet
[ʤæd]	jad	['ʤædɪf]	jadif
[ʤæd]	jaf	[ʤa:'fɪn]	jafen
[ʤeg]	jeg	['ʤegɪn]	jegin
[ʤev]	jev	[ʤeɪ'væt]	jevat
[ʤʌf]	juf	[ʤə'fɪs]	jufis
[ʤʌk]	juk	['ʤu:kɪl]	jukel
[kæm]	kam	[kæ'mʌt]	kamut
[keb]	keb	['kebæm]	kebam
[lær]	lar	[la: 'rɒf]	larof
[lek]	lek	['lekɪv]	lekiv
[lep]	lep	[lə'pɒd]	lepod
[lɪg]	lig	[lɪ'gɒp]	ligop
[lɒf]	lof	[ləʊ'fʌz]	lofuz
[lʌd]	lud	['lu:dem]	ludem
[lʌm]	lum	['ləmɒt]	lumot
[mʌs]	mus	['mju:sɒv]	musov

monosyllabic CVC words		bisyllabic CVCVC words	
[nær]	nar	[na: 'rel]	narel
[næt]	nat	['na:təf]	natuf
[nɪs]	nis	['nɪsʊr]	nisur
[nɒs]	nos	['nɒseg]	noseg
[nʌg]	nug	['nəgʌm]	nugom
[nʌp]	nup	['nu:pæf]	nupaf
[pes]	pes	['pesɛ:r]	pesir
[pɪf]	pif	['pɪfa:r]	pifar
[pɪər]	pir	['pɪrʌk]	piruk
[pɒf]	pof	['pɒfʌs]	pofus
[rʌd]	rud	[rə'dɪl]	rudil
[tæf]	taf	['tæfɪp]	tafep
[teg]	teg	['tegɒr]	tegor
[væm]	vam	['vamɪks]	vamex
[vek]	vek	['vekəs]	vekas
[vɒg]	vog	['vəʊgʌd]	vogud
[vɒk]	vok	['vɒkæz]	vokaz
[vʌg]	vug	['vu:gæp]	vugap
[wem]	wem	['wemɪk]	wemic
[zek]	zec	['zekɪd]	zecid
[zel]	zel	[ze'lɒn]	zelon
[zep]	zep	['zɛpæk]	zepak
[zɪn]	zin	[zɪ'nev]	zinev
[zɒg]	zog	['zɒ:gɪk]	zogik
[zɒm]	zom	['zɒ:mən]	zomun

Table 3: Foils used for 2AFC test during the Familiarisation phase in Experiment 6, and CVC words used as a basis.

	bok		min
	buf		mof
	cal		nop
	ceb		peg
	cur		rak
	dem		sap

	fal		sif
	fun		tel
	gav		ten
	jah		tim
	kil		ver
	kor		zoh

 kul	 zup
 lid	 min
 lon	 mof

Appendix B

ANOVA tables

Table 1: Results of the 2x2x2 ANOVA for Experiment 1: (1) = Familiarity (familiar vs. unfamiliar X words), (2) = Grammaticality (grammatical vs. ungrammatical strings), (3) = Condition (+Gestalt vs. -Gestalt).

Source	**df Effect**	**df Error**	**MS Error**	**F**	**Sig.**	**Partial Eta Squared**	**Observed Power**
1	1	54	233.49	48.54	***	.47	1.0
2	1	54	1341.43	28.49	***	.35	1.0
3	1	54	486.04	2.02	n.s.	.04	.29
12	1	54	41.33	.04	n.s.	.001	.05
13	1	54	233.49	.30	n.s.	.005	.08
23	1	54	1341.43	.85	n.s.	.02	.15
123	1	54	41.33	.89	n.s.	.02	.15

Note:
n.s. $p > .05$
* $p < .05$
** $p < .01$
*** $p < .001$

Table 2: Results of the 2x2x2 ANOVA for Experiment 2: (1) = Familiarity (familiar vs. unfamiliar X chunks), (2) = Grammaticality (grammatical vs. ungrammatical strings), (3) = Condition (+Gestalt vs. -Gestalt).

Source	df Effect	df Error	MS Error	F	Sig.	Partial Eta Squared	Observed Power
1	1	46	264.65	44.12	***	.49	1.0
2	1	46	907.97	22.12	***	.33	1.0
3	1	46	575.54	.50	n.s.	.01	.11
12	1	46	38.45	2.94	n.s	.06	.39
13	1	46	264.65	.06	n.s.	.001	.06
23	1	46	907.97	9.82	**	.18	.87
123	1	46	38.45	.38	n.s.	.01	.09

Note:
n.s. p > .05
* p < .05
** p < .01
*** p < .001

Table 3: Results form the 3x2 ANOVA for Experiment 2: (1) = Grammaticality (grammatical vs. ungrammatical strings), (2) = Stimuli type (-Gestalt, One-word, Nine-letter).

Source	df Effect	df Error	MS Error	F	Sig.	Partial Eta Squared	Observed Power
1	1	77	507.46	33.87	***	.31	1
2	2	77	229.84	1.1	n.s.	.03	.23
12	2	77	507.46	4.41	*	.10	.74

Note:
n.s. p > .05
* p < .05
** p < .01
*** p < .001

Table 4: Results from the 2x2x4 ANOVA for Experiment 3: (1) = Familiarity (familiar vs. unfamiliar X chunks), (2) = Grammaticality (grammatical vs. ungrammatical strings), (3) = Condition (–Gestalt, vs. Shape vs. Holistic vs. Componential).

Source	**df Effect**	**df Error**	**MS Error**	**F**	**Sig.**	**Partial Eta Squared**	**Observed Power**
1	1	116	434.60	107.43	***	.48	1.0
2	1	116	986.98	25.82	***	.18	1.0
3	3	116	367.18	.42	n.s.	.01	.13
12	1	116	78.80	5.72	*	.05	.66
13	3	116	434.60	5.37	**	.12	.93
23	3	116	986.98	4.07	**	.10	.83
123	3	116	78.80	1.98	n.s.	.05	.50

Note:
n.s. $p > .05$
* $p < .05$
** $p < .01$
*** $p < .001$

Table 5: Results of the 2x2x4x2 ANOVA for Experiment 4: (1) = Familiarity (familiar vs. unfamiliar X chunks), (2) = Grammaticality (grammatical vs. ungrammatical strings), (3) = Condition (–Gestalt, vs. Shape vs. Holistic vs. Componential), (4) = Language (L1a vs. L2a).

Source	df Effect	df Error	MS Error	F	Sig.	Partial Eta Squared	Observed Power
1	1	88	424.71	25.84	***	.23	1.0
2	1	88	1868.59	56.74	***	.39	1.0
3	3	88	334.72	.45	n.s.	.02	.14
4	1	88	334.72	.40	n.s.	.00	.10
12	1	88	48.55	.22	n.s.	.00	.08
13	3	88	424.71	1.82	n.s.	.06	.46
23	3	88	1868.59	5.35	**	.15	.92
14	1	88	424.71	1.05	n.s.	.01	.17
24	1	88	1868.59	.71	n.s.	.01	.13
34	3	88	334.72	.31	n.s.	.01	.11
123	3	88	48.55	.42	n.s.	.01	.13
124	1	88	48.55	.03	n.s.	.00	.05
134	3	88	424.71	1.11	n.s.	.04	.29
234	3	88	1868.59	1.47	n.s.	.05	.38
1234	3	88	48.55	.53	n.s.	.02	.15

Note:

n.s. $p > .05$
* $p < .05$
** $p < .01$
*** $p < .001$

Table 6: Results of the cross-experimental analysis in Chapter 4: (1) = Presentation Mode (simultaneous vs. sequential), (2) = Grammaticality (grammatical vs. ungrammatical strings), (3) = Stimulus Type (+Gestalt vs. –Gestalt vs. Shape vs. Componential).

Source	**df Effect**	**df Error**	**MS Error**	**F**	**Sig.**	**Partial Eta Squared**	**Observed Power**
1	3	200	208.10	.48	n.s.	.00	.11
2	1	200	748.00	97.12	***	.33	1.0
3	2	200	208.10	.69	n.s.	.01	.19
12	1	200	748.00	5.96	*	.03	.68
13	3	200	208.10	.97	n.s.	.01	.26
23	3	200	748.00	9.27	***	.12	1.0
123	3	200	748.00	1.05	n.s.	.02	.28

Note:
n.s. p > .05
* p < .05
** p < .01
*** p < .001

Table 7: Results of the ANOVA for Experiment 5: (1) = Familiarity (familiar vs. unfamiliar X item), (2) = Grammaticality (grammatical vs. ungrammatical strings), (3) = Condition (+Gestalt vs. –Gestalt vs. Shape vs. Componential), (4) = Language (L1 vs. L2).

Source	**df Effect**	**df Error**	**MS Error**	**F**	**Sig.**	**Partial Eta Squared**	**Observed Power**
1	1	88	209.21	44.94	***	.34	1.0
2	1	88	889.64	21.87	***	.20	1.0
3	3	88	420.22	3.51	*	.11	.76
4	1	88	420.22	4.96	*	.05	.60
12	1	88	52.17	.08	n.s.	.00	.06
13	3	88	209.21	7.03	***	.19	.98
23	3	88	889.64	3.46	*	.11	.76
14	1	88	209.21	.06	n.s.	.00	.06
24	1	88	889.64	.07	n.s.	.00	.06
34	3	88	420.22	1.48	n.s.	.05	.38
123	3	88	52.17	1.15	n.s.	.04	.30
124	1	88	52.17	.01	n.s.	.00	.05
134	3	88	209.21	.92	n.s.	.03	.24
234	3	88	889.64	.74	n.s.	.03	.20
1234	3	88	52.17	1.1	n.s.	.04	.29

Note:
n.s. $p > .05$
* $p < .05$
** $p < .01$
*** $p < .001$

Table 8: Results of cross-modal analysis in Chapter 5: (1) = Modality (auditory vs. visual), (2) = Grammaticality (grammatical vs. ungrammatical strings), (3) = Condition (+Gestalt vs. –Gestalt vs. Shape vs. Componential).

Source	**df Effect**	**df Error**	**MS Error**	**F**	**Sig.**	**Partial Eta Squared**	**Observed Power**
1	1	184	192.45	4.34	*	.02	.55
2	1	184	691.10	78.25	***	.30	1.0
3	3	184	192.45	2.43	n.s.	.04	.60
12	1	184	691.10	12.53	**	.06	.94
13	3	184	192.45	1.79	n.s.	.03	.46
23	3	184	691.10	5.47	**	.08	.94
123	3	184	691.10	3.99	**	.06	.83

Note:
n.s. p > .05
* p < .05
** p < .01
*** p < .001

Tables 9 – 10 for cross-modal analysis for Chapter 5:

Table 9: Results of +Gestalt analysis: (1) = Modality (visual vs. Auditory), 2 = Grammaticality (grammatical vs. ungrammatical)

Source	**df Effect**	**df Error**	**MS Error**	**F**	**Sig.**	**Partial Eta Squared**	**Observed Power**
1	1	46	179.98	.26	n.s.	.01	.08
2	1	46	965.97	21.05	***	.31	.99
12	1	46	965.97	2.91	n.s.	.06	.39

Table 10: Results of -Gestalt analysis: (1) = Modality (visual vs. Auditory), 2 = Grammaticality (grammatical vs. ungrammatical)

Source	**df Effect**	**df Error**	**MS Error**	**F**	**Sig.**	**Partial Eta Squared**	**Observed Power**
1	1	46	163.24	.31	n.s.	.01	.08
2	1	46	994.38	19.09	***	.29	.99
12	1	46	9994.38	.00	n.s.	.00	.05

Table 11: Results of Shape analysis: (1) = Modality (visual vs. Auditory), 2 = Grammaticality (grammatical vs. ungrammatical)

Source	**df Effect**	**df Error**	**MS Error**	**F**	**Sig.**	**Partial Eta Squared**	**Observed Power**
1	1	46	198.08	6.34	*	.12	.69
2	1	46	644.22	39.55	***	.46	1.0
12	1	46	644.22	21.59	***	.32	1.0

Table 12: Results of Componential analysis: (1) = Modality (visual vs. Auditory), 2 = Grammaticality (grammatical vs. ungrammatical)

Source	df Effect	df Error	MS Error	F	Sig.	Partial Eta Squared	Observed Power
1	1	46	228.51	2.28	n.s.	.05	.32
2	1	46	159.84	3.94	n.s.	.08	.49
12	1	46	159.84	1.33	n.s.	.03	.21

Note:
n.s. p > .05
* p < .05
** p < .01
*** p < .001

Table 13: Results of ANOVA for AGL task in Experiment 6: (1) = Familiarity (familiar vs. novel X item), (2) = Grammaticality (grammatical vs. ungrammatical strings), (3) = Training (0-block vs. 1-block vs. 4-block vs. 8+2-block).

Source	**df Effect**	**df Error**	**MS Error**	**F**	**Sig.**	**Partial Eta Squared**	**Observed Power**
1	1	76	281.65	30.52	***	.27	1.0
2	1	76	270.05	1.12	n.s.	.01	.18
3	3	76	401.09	2.22	n.s.	.08	.54
12	1	76	58.53	4.30	*	.05	.54
13	3	76	281.65	1.94	n.s.	.07	.48
23	3	76	270.05	.08	n.s.	.00	.06
123	3	76	58.53	.68	n.s.	.03	.19

Note:
n.s. p > .05
* p < .05
** p < .01
*** p < .001

References

Abla, D., & Okanoya, K. (2009). Visual statistical learning of shape sequences: An ERP study. *Neuroscience Research, 64 (2)*, 185 – 190.

Aiello, L.C., & Dunbar, R.I.M. (1993). Neocortex size, group size and the evolution of language. *Current Anthropology*. *34*, 184 – 193.

Atkinson, A.P., & Wheeler, M. (2004). The grain of domains: The evolutionary-psychological case against domain-general cognition. *Mind and Language, 19 (2)*, 147 – 176.

Bahlmann, J., & Friederici, A.D. (2006). FMRI investigations of the processing of simple linear and embedded hierarchical structures: An artificial grammar task. *Journal of Cognitive Neuroscience, Supplement*.

Bahrick, L.E. & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology, 36*, 190 – 201.

Bahrick, L.E., Lickliter, R., Castellanos, I., & Vaillant-Molina, M. (2010). Increasing task difficulty enhances effects of intersensory redundancy: testing a new prediction of the intersensory redundancy hypothesis. *Developmental Science, 13(5)*, 731 – 737.

Baldwin, K.B., & Kutas, M. (1997). An ERP analysis of implicit structured sequences learning. *Psychophysiology, 34*, 74 – 86.

Bates, E.A.(1994). Modularity, domain specificity and the development of language. *Discussions in Neuroscience, 10*, 136 – 149.

Beaman, C.P. (2002). Inverting the modality effect in serial recall. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 55A*, 371-389.

Brooks, L.R., & Vokey, J.R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al. (1989). *Journal of Experimental Psychology: General, 120*, 316 – 323.

Butterworth, B. (1980). Evidence from pauses in speech. In: B. Butterworth (Ed.), *Language production, Volume 1: Speech and Talk* (pp.155 – 176). London: Academic Press.

Carey, S., & Spelke, E. (1994). Domain-specific knowledge and conceptual change. In L.A. Hirschfeld & S.A. Gelman (Eds.), *Mapping the mind: Domain* *specificity in cognition and culture* (pp. 169 – 200). New York: Cambridge University Press.

Carruthers, P. (2006). The case for massively modular models of mind. In R. Stainton (Ed.), *Contemporary debates in cognitive science* (pp. 3 – 21). Oxford, England: Blackwell.

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

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

Chomsky, N. (1980). On cognitive structures and their development: A reply to Piaget. In M. Piatelli-Palmarini (Ed.), *Language and learning: The debate between Jean Piaget and Noam Chomsky* (pp. 142 - 162). Cambridge, MA: Harvard University Press.

Chomsky, N. (1986). *Knowledge of language: It's nature, origin and use*. Westport, Connecticut, London: Praeger.

Chomsky, N. (1988). *Language and problems of knowledge: The Managua lectures*. Cambridge, MA: MIT Press.

Christiansen, M.H., & Chater, N. Language as shaped by the brain. *Behavioral and Brain Sciences*, 31, 489 – 558.

Christiansen, M.H., Conway, C.M., & Onnis, L. (2007). Neural responses to structural incongruencies in language and statistical learning point to a similar underlying mechanism. In D.S. McNamara & J.G. Trafton (Eds.), *Proceedings of the 29th Annual Conference of the Cognitive Science Society* (pp. 173 – 178). Austin, TX: Cognitive Science Society.

Christiansen, M.H., & Curtin, S.L. (1999). The power of statistical learning: No need for algebraic rules. In *The Proceedings of the 21st Annual Conference of the Cognitive Science Society* (pp. 114-119). Mahwah, NJ: Lawrence Erlbaum.

Clark, H. H., & Clark, E. V. (1977). *Psychology and language: An introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.

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

Conway, C., & Christiansen, M.H. (2006). Statistical learning within and between modalities: Pitting abstract knowledge against stimulus specific representations. *Psychological Science, 17*, 905 – 912.

Conway, C., & Christiansen, M.H. (2009). Seeing and hearing in space and time: Effects of modality and presentation rate on implicit statistical learning. *European Journal of Cognitive Psychology, 21*, 561 – 580.

Conway, C.M.. & Pisoni, D.B. (2008). Neurocognitive basis of implicit learning of sequential structure and its relation to language processing. *Annals of the New York Academy of Sciences*, *1145*,113-131.

Corballis, M.C (2003). From hand to mouth: The gestural origins of language. In M.H. Christiansen & S. Kirby (Eds.), *Language Evolution* (pp. 201 – 218). Oxford, New York: Oxford Univeristy Press.

Cosmides, L., & Tooby, J. (1994). Origins of domain-specificity: The evolution of functional organization. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain-specificity in cognition and culture* (pp.85 – 116). New York: Cambridge University Press.

Cosmides, L., & Tooby, J. (2006). Evolutionary psychology, moral heuristics, and the law. In G. Gigerenzer & C. Engel (Eds.), *Heuristics and the Law* (pp.181 - 212). Cambridge, MA: MIT Press.

Creel, S.C., Newport, E.L., & Aslin, R.A. (2004). Distant melodies: Statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30 (5)*, 1119 – 1130.

DeCasper, J.A., & Spence, M.J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development, 9 (2)*, 133 – 150.

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

Dienes, Z., Altmann, G.T.M., & Gao, S.-J. (1999). Mapping across domains without feedback: A neural network model of transfer of implicit knowledge. *Cognitive Science, 23,* 53 – 82.

Dunbar, R.I.M (2003). The origin and subsequent evolution of language. In M.H. Christiansen & . Kirby (Eds.), *Language Evolution* (pp.219 – 234), Oxford, New York: Oxford Univeristy Press.

Elman, J.L., Bates, E.A., Johnson, M.H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Evans, N., & Levinson, S.C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences, 32*, 429 – 492.

Farmer, T.A., Christiansen, M.H., & Monaghan, P. (2006). Phonological typicality influences on-line sentence comprehension. *Proceedings of the National Academy of Sciences, 103,* 12203 – 12208.

Field, D.J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America*, *4*, 2379 – 2394.

Fiser, J., & Aslin, R. (2002a). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *28*, 458 – 467.

Fiser, J., & Aslin, R. (2002b). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences, 99,* 15822 – 15826.

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

Fodor, J.A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge: MIT Press.

Frank, M.C., Slemmer, J.A., Marcus, G.F., & Johnson, S.P. (2009). Information from multiple modalities helps 5-month-olds learn abstract rules. *Developmental Science*, *12, (4)*, 504-509.

Friederici, A.D. (2004). Processing local transitions versus long-distance syntactic hierarchies. *Trends in Cognitive Sciences, 8 (6)*, 245 – 247.

Friederici, A.D., Bahlmann, J., Heim, S., Schubotz, R.I., & Anwander, A. (2006). The brain differentiates human and nonhuman grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences*, *103*, 2458 – 2463.

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

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

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

Glenberg, A.M., & Fernandez, A. (1988). Evidence for auditory temporal distinctiveness: Modality effects in order and frequency judgments. *Journal of Experimental Psychology, 14*, 728-739.

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

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

Gómez, R.L., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy, 7 (2)*, 183 – 206.

Gómez, R.L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences, 4 (5)*, 178 – 186.

Gómez, R.L., Gerken, L., & Schvaneveldt, R.W. (2000). The basis of transfer in artificial grammar learning. *Memory and Cognition*, *28 (2)*, 253 – 263.

Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behaviour*. Cambridge, Massachusetts: Harvard University Press.

Goschke, T. (1998). Implicit learning of perceptual and motor sequences: evidence for independent learning systems. In M.A. Stadler & P.A. Frensch (Eds.), *Handbook of implicit learning* (pp.410-444). London: Sage Publications.

Hagoort, P., Brown, C.M., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, *8,* 439 – 484.

Hahne, A., & Friederici, A.D. (1999). Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11*, 193 – 204.

Hauser, M.D., Chomsky, N., & Fitch, W.T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science, 298*, 1569 – 1579.

Hauser, M.D., Newport, E.L., & Aslin, R.N. (2001). Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition, 78*, 53 – 64.

Henderson, J.M. (2007). Regarding scenes. *Current Directions in Psychological Science*, *16(4)*, 219 – 222.

Henderson, J.M. & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology, 50*, 243 – 71.

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

Höhle, B., Schmitz, M., Santelmann, L.M., & Weissenborn, J. (2006). The recognition of discontinuous verbal dependencies by German 19-month-olds: Evidence for lexical and structural influences on children's early processing capacities. *Language Learning and Development, 2 (4)*, 277 – 300.

Hollingworth, A., & Henderson, J.M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 113 – 136.

Johansson, T. (2009). Strengthening the case for stimulus-specificity in artificial grammar learning: No evidence for abstract representations with extended exposure. *Experimental Psychology, 56 (3)*, 188-197.

Johnson, E.K., & Jusczyk, P.W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, 44 (4), 548-567.

Johnson, S.P., Fernandes, K.J., Frank, M.C., Kirkham, N.Z., Marcus, G.F., Rabagliati, H., & Slemmer, J.A. (2009). Abstract rule learning for visual sequences in 8-and 11-month-olds. *Infancy*, *14*, 2 - 18.

Jusczyk, P.W. (1999). How infants begin to extract words from speech. *Trends in Cognitive Sciences, 3 (9)*, 323 – 328.

Jusczyk, P.W., Houston, D.M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*, 159 – 207.

Kemler Nelson, D.G., Jusczyk P.W., Mandel, D.R., Myers, J., Turk, A. & Gerken, L.A. (1995). The headturn preference procedure for testing auditory perception. *Infant Behavior and Development*, *18*, 111 – 116.

Khalidi, M.A.(2001). Innateness and domain specificity. *Philosophical Studies, 105,* 191 – 210.

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

Kircher,T.T.J, Brammer, M.J., Levelt, W., Bartels, M., & McGuire, P.K. (2004). Pausing for thought: engagement of left temporal cortex during pauses in speech. *Neuroimage, 21*, 84-90.

Kirkham, N.Z., Slemmer, J.A., & Johnson, S.P. (2002). Visual statistical learning in infancy: Evidence for a domain-general learning mechanism. *Cognition, 83*, B35 – B42.

Kirkham, N.Z., Slemmer, J.A., Richardson, D.C., & Johnson, S.P. (2007). Location, location, location: Development of spatiotemporal sequence learning in infancy. *Child Development, 78*, 1559 – 1571.

Kuhn, G. & Dienes, Z. (2005). Implicit learning of nonlocal musical rules: Implicitly learning more than chunks. *Journal of Experimental Psychology: Learning, Memory and Cognition, 31 (6)*, 1417 – 1432.

Lau, E., Stroud, C., Plesch, S., & Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and Language, 98*, 74 – 88.

Leech, R., Holt, L.L., Devlin, J.T., & Dick, F. (2009). Expertise with artificial nonspeech sounds recruits speech-sensitive cortical regions. *The Journal of Neuroscience, 29(16)*, 5234 – 5239.

Lewkowicz, D. (2000). The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychological Bulletin, 126*, 281–308.

Liversedge, S.P., & Findlay, J.M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Science, 4(1)*, 6 – 14.

MacGregor, Corley and Donaldson (forthcoming). Listening to the sound of silence: Investigating the consequences of disfluent silent pauses in speech for listeners.

Marcus, G.F., Fernandes, K., & Johnson, S. (2007). Infant rule-learning facilitated by speech, *Psychological Science, 18*, 387 – 391.

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

Mathews, R.C., & Roussel, L.G. (1997). Abstractness of implicit knowledge: A cognitive evolutionary perspective. In D.C. Berry (Ed.), How *implicit is implicit learning?* (pp. 13 – 47). New York: Oxford University Press.

Mattys, S.L., & Jusczyk, P.W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, 78 (2), 91-121.

Mattys, S.L., Jusczyk, P.W., Luce, P.A., & Morgan, J.L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, 38, 465-494.

Misyak, J.B., Christiansen, M.H., & Tomblin, J.B. (2010). On-line individual differences in statistical learning predict language processing. *Frontiers in Psychology*, 1, 1 – 9.

Mueller, J.L., Bahlmann, J., & Friederici, A.F. (2008). The role of pause cues in language learning: The emergence of event-related potentials related to sequence processing. *Journal of Cognitive Neuroscience, 20 (5)*, 892-905.

Newport, E.L., & Aslin, R.N. (2004). Learning at a distance I. Statistical learning of nonadjacent dependencies. *Cognitive Psychology, 48*, 127 – 162.

Newport, E.L., Hauser, M.D., Spaepen, G., Aslin, R.N. (2004). Learning at a distance II. Statistical learning of nonadjacent dependencies in a non-human primate. *Cognitive Psychology, 49*, 85 – 117.

Nuthman, A., & Henderson, J.M. (2010). Object-based attentional selection in scene viewing. *Journal of Vision, 10(8)*, 1 – 19.

O'Grady, W., Dobrovolsky, M., & Katamba, F. (Eds.). (1996). *Contemporary Linguistics: An Introduction*. Harlow: Longman.

Ohms, V.R., Gill, A., Van Heijningen, C.A.A., Beckers, G.J.L., & Ten Cate, C. (2010). Zebra finches exhibit speaker-independent phonetic perception of human speech. *Proceedings of the Royal Society, 277*, 1003 – 1009.

Olaha, J.J. (1994). Toward a universal, phonetically-based, theory of vowel harmony. *Proceedings of the 3rd International Conference on Spoken Language Processing* (pp. 491 – 494).Yokohama, Japan: IEEE Press.

Onnis, L., Christiansen, M.H., Chater, N., & Gómez, R.L. (2003). Reduction of uncertainty in human sequential learning: Evidence form artificial language learning. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the 25th Annual*

Conference of the Cognitive Science Society (pp.886 – 891). Boston, MA: Cognitive Science Society.

Onnis, L., Monaghan, P., Christiansen, M.H., & Chater, N. (2004). Variability is the spice of learning, and a crucial ingredient for detecting and generalizing in nonadjacent dependencies. In K. Forbus, D. Gentner & T. Regier (Eds.), *Proceedings of the 26th Annual Conference of the Cognitive Science Society* (pp. 1047 – 1052). Chicago, Illinois: Cognitive Science Society.

Onnis, L., Monaghan, P., Richmond, K., & Chater, N. (2005). Phonology impacts segmentation in speech processing. *Journal of Memory and Language, 53*, 225 – 237.

Pearce, M.T., & Wiggins, G.A. (2006). Expectation in melody: The influence of context and learning. *Music Perception, 23 (5),* 377 – 405.

Peña, M., Bonatti, L.L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science, 298*, 604 – 607.

Perfors, A., Tenenbaum, J.B., & Regier, T. (forthcoming). The learnability of abstract syntactic principles. *Cognition*.

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

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences, 13 (4)*, 707 – 784.

Prinz, J.J. (2006). Is the mind really modular? In R. Stainton (Ed.), *Contemporary debates in cognitive science* (pp. 22 – 36). Oxford, England: Blackwell.

Pullum, G.K., & Scholz, B.C. (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review, 19*, 9 – 50.

Pulvermüller, F. (2010). Brain embodiment of syntax and grammar: Discrete combinatorial mechanisms spelt out in neuronal circuits. *Brain and Language*, *112*, 167 – 179.

Quinn, P.C., Burke, S., & Rush, A. (1993). Part-whole perception in early infancy: Evidence for perceptual grouping produced by lightness similarity. *Infant Behavior and Development, 16 (1)*, 19 – 42.

Radford, A. (1997). *Syntactic theory and the structure of English: A minimalist approach*. Cambridge: Cambridge University Press.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC nonword databse. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 55(A)*, 1339 – 1362.

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

Reichle, E.D., Reyner, K., & Pollatesk, A. (2003). The E – Z reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*, 445 – 526.

Saffran, J.R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, *47*, 172 – 196.

Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926 – 1928.

Saffran, J.R., Johnson, E.K., Aslin, R.N., & Newport, E.L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27 – 52.

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

Saffran, J.R., Pollack, S.D., Seibel, R.L., & Shkolnik, A. (2007). Dog is a dog is a dog: Infant rule learning is not specific to language. *Cognition*, *105*, 669 – 680.

Santelmann, L.M., & Jusczyk, P.W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition, 69*, 105 – 134.

Schellenberg, E.G., Adachi, M., Purdy, K.T., & McKinnon, M.C. (2002). Expectancy in melody: Tests of children and adults. *Journal of Experimental Psychology: General, 131 (4)*, 511 – 537.

Smith, K., Brighton, H., & Kirby, S. (2003). Complex systems in language evolution: the cultural emergence of compositional structure. *Advances in Complex Systems, 6*, 537 – 558.

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

Spelke, E.S., & Kinzler, K.D. (2007). Core knowledge. *Developmental Science, 10,* 89 – 96.

Spence, M.J., & DeCasper, A.J. (1987). Prenatal experience with low-frequency maternal-voice sounds influence neonatal perception of maternal voice samples. *Infant Behavior and Development, 10 (2),* 133 – 142.

Suomi, K., McQueen, J.M., & Cutler, A. (1997). Vowel harmony and speech segmentation in Finnish. *Journal of Memory and Language, 36*, 422 – 444.

Takashi, M., Yamada, H., & Okanoya, K. (2010). Statistical and prosodic cues for song segmentation learning by Bengalese finches (*Lonchura striata* var. d*omestica*). *Ethology, 116,* 481 – 489.

Tomasello, M. (2008). *Origins of Human Communication*. Cambridge, MA: MIT Press.

Tomasello, M., Strosberg, R., & Akhtar, N. (1996). Eighteen-month-old children learn words in non-ostensive contexts. *Journal of Child Language,23,*157 - 17.

Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp.19 – 136). New York: Oxford University Press.

Tunney, R.J., & Altmann, G.T.M. (1999). The transfer effect in artificial grammar learning: Reappraising the evidence on the transfer of sequential dependencies. Journal of Experimental Psychology: *Learning, Memory and Cognition, 25 (5),* 1322 – 1333.

Turk-Browne, N.B., Jungé, J., & Scholl, B.J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General, 134 (4)*, 552 – 564.

Wade, T. & Holt, L.L. (2005). Incidental categorization of spectrally complex non-invariant auditory stimuli in a computer game task. *The Journal of the Acoustical Society of America, 118*, 2618 – 2633.

Wertheimer, M. (1938). Laws of organization in perceptual forms. In W. Ellis (Ed. And Trans.), *A source book of Gestalt psychology*. London: Routledge & Kegan Paul.

Wertheimer, M. (1944). Gestalt theory. *Social Research, 11*, 78 – 99.

Zellner, B. (1994). Temporal structure of speech. In E. Keller (Ed.), *Fundamentals of speech synthesis and speech recognition* (pp. 41-62). Chichester: John Wiley.