8-28-2018 11:00 AM

Exploring the Process of Statistical Language Learning

Nicolette B. Noonan
The University of Western Ontario

Supervisor
Archibald, Lisa M. D.
The University of Western Ontario Joint Supervisor
Joanisse, Marc F.
The University of Western Ontario

Graduate Program in Psychology
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

Abstract

In this thesis, I investigated the process of statistical word segmentation using a combination of behavioural, clinical, and neuroimaging approaches. Prior research has largely focused on the outcome of statistical learning approaches, with little research attention paid to the process of learning. In this body of research, I sought to address this issue. In Chapter 2, I examined how domain-specific and domain-general working memory interference effects on a statistical word segmentation task. I found that when completing a concurrent visuospatial or verbal working memory task, statistical language learning was impaired. Thus, this study provided some evidence that domain-general working memory may support statistical language learning. In Chapter 3, I further investigated how cognitive processes, including language and working memory, are involved in statistical learning across domains. In this study, school-aged children with and without a developmental language disorder (DLD) completed a statistical language learning task and a visual statistical learning task. I found that those with DLD did not differ from typically developing children on either statistical learning task, and that performance across groups was meager for the statistical language learning task, and not above chance levels for the visual statistical learning task. Further, performance on the statistical learning tasks was not associated with other cognitive processes. This raised the possibility that an alternative measurement approach may be better suited to examine statistical learning. I addressed this issue in Chapter 4, where I measured event related potentials (ERPs) using electroencephalography (EEG) during exposure to a structured, unsegmented language. I found that statistical learning performance was related to neural responses to the structured linguistic input, and that ERPs were modulated as a function of language exposure, revealing the dynamic nature of statistical learning. Chapter 5 discusses the relevant findings of this thesis in relation to the current state of affairs in statistical learning research, and presents recommendations for future research in examining the process of statistical learning.

Keywords

Statistical language learning, language acquisition, implicit learning, working memory, developmental language disorder, specific language impairment, event-related potentials

Acknowledgements

I am incredibly grateful for the support I received from my friends, family, and advisors in completing this thesis.

To Dr. Lisa Archibald: You took a chance on me seven years ago. You showed me not only how far I could go, but guided and supported me at every step along the way. Your mentorship and your dedication to your research continues to be a model to me.

To Dr. Marc Joanisse: I walked into your classroom in eight years ago, not knowing that those short four months would eventually lead me onto this long path. I'm a better scientist because of the example you have set. Your guidance and support as a mentor has been truly invaluable.

I would also like to thank the other scientists I have worked with along the way. Thank you to my advisory committee, Dr. Debra Jared and Dr. Ryan Stevenson, for your helpful comments, input, and support. Thank you to the Developmental Psychology Research Group and the Language and Concepts Research Group, or as they are now known known, the Cognitive, Developmental, and Brain Sciences Research Group, for your useful feedback throughout six brownbag presentations. I would be remiss to not acknowledged the Brain and Mind Institute. The experience of being part of this scientifically rigorous research group is truly unparalleled.

To my lab mates past and present: Alex Cross, Christina Van den Bosch der Nederlander, Christine Moreau, Joseph Nidel, Leah Brainin, Emily Nichols, Veronica Whitford, Meghan Vollebregt, Alyssa Kuiack, Theresa Pham, Laura Pauls, Areej Balilah, Monica DaSilva, Allison Partridge, and all the RAs and undergraduate students who joined in along the way. Thank you all for your friendship, and for your academic and emotional support throughout the years. It has been a pleasure to work with both the LWM and LRCN lab groups. I'll miss our elevenses and 12pm-sharp lunch breaks.

To my family: Thank you for continually supporting and encouraging me. I would like to extend a special thank you to my grandparents, without whom, none of this would have been possible.

And finally, to Kyle: Thank you for being a supportive partner, and for patiently listening to me while I described every minute detail of this thesis to (/at) you.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1: Introduction

The process of human language acquisition is an extraordinary cognitive phenomenon. Infants are born without the ability to *use* language, but remarkably, they are well underway in developing the ability to *learn* language. What is perhaps most remarkable is the rapidity with which infants learn the unique characteristics of their native language. For instance, at birth, infants can discriminate languages with different prosodic structure (Mehler, Jusczyk, Lambertz, Halsted, Bertoncini, & Amiel-Tison, 1988), with the ability to discriminate between their native language and a language in the same rhythmic class developing around 4-5 months of age (Nazzi & Ramus, 2003). By 7 months of age, infants show a preference for the stress, melody, and phonotactics of their native language (Jusczyk, 2002), and have become attuned to the phonetic categories within their native language (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992). By their first birthday, infants begin to produce their first words, and begin using full sentences by 3 years of age. And, by 4-5 years of age, young children have mastered nearly all the rules of their native language, and can both use and understand most linguistic structures. Given how quickly infants seem to master many of the fundamentals of their native language, it seems only natural that exploring this phenomenon has been of great interest to psychologists, linguists, and neuroscientists. My research seeks to examine one way in which humans acquire language, namely, the process of statistical learning.

1.1 Early Theories on Language Acquisition

Early theories of language acquisition sought an explanation for the rapidity with which infants acquired their native language. The most prominent theories from Piaget (*see* Piaget, 2005), Skinner (1957) and Chomsky (e.g., Chomsky, 1965) varied substantially in how they viewed the interaction between general cognition and language faculties. Fitting with his view on cognitive development as a movement through distinct stages, Piaget saw language acquisition as a similar stage-like phenomenon (*see* Piaget, 2005). In his view, a child's language developed from being largely egocentric, to being symbolic and reflective of socialized knowledge and behaviour. Additionally, Piaget saw language development as being constrained by domain-neural cognitive processes. Similar to Piaget, Skinner (1957) argued that language developed like other cognitive phenomena, though his theory was constrained within his own views on behaviorism. Skinner speculated that language was acquired via a process of operant conditioning, that is, language acquisition was due to an organism's history of reinforcement and shaping. This view was strictly a learning-based account, and viewed the language-learner as a *tabula rasa* or blank slate.

Chomsky's (1959) proposals on language acquisition were a direct critique of Skinner's (1957) *tabula rasa* view. Chomsky (1959) argued that reinforcement learning has little to do with language acquisition as instead, he saw the language learner as an "innate grammarian". Central to his argument was his notion of the "poverty of the stimulus", as the language learner receives too little input to account for their language outcomes. Because of this impoverished input, Chomsky suggested that the formal structures of language must be innately specified. This nativist view of an innate specification for language was echoed by others, including Eimas, Siqueland, Juszcky and Vigorito (1971), who also concluded that the ability to acquire a language must be part of a human's biological make-up. These built-in constraints, it was thought, could only be learned by a specialized "language acquisition device" (Chomsky, 1965; Gleitman &

Wanner, 1982; Pinker, 1984). In a similar argument, Pinker (1994) suggested that a "language instinct" was based on micro-circuitry that detailed the innate knowledge of the principles of grammar, and was determined by natural selection (*also see* Pinker & Bloom, 1990). At the same time that Chomsky (1965) was arguing for a specialized and innately specified "language acquisition device", Lenneberg (1967) was advancing his proposal that language is a system deeply constrained by biology. In his influential book, *Biological Foundations of Language*, Lenneberg (1967) argued for a "critical period" for language, based on observations that optimal language development can only be acquired through birth until the onset of puberty (*see* Werker & Tees, 2005 *for a discussion*). Lenneberg's (1967) proposal was found to be consistent with later findings related to second language acquisition throughout development (e.g., Johnson & Newport, 1989). Although there are theoretical distinctions between many of these nativist claims, this body of work has converged at the conclusion that language must be constrained by our biological makeup.

These nativist claims have been, and continue to be, strongly influential theories of language acquisition. However, their validity has been questioned by a growing body of research indicating the importance of interactions with the environment during the learning process. Beginning in the 1980s, research on infants' phonetic perception challenged nativist views, as this research showed that young infants could engage in a detailed mapping between the distributional properties of environmental input to a phonetic percept (Kuhl et al., 1992; Werker & Tees, 1980; *see* Kuhl, 2000), highlighting how experience-dependent learning can operate within some of the earliest stages of language acquisition. Moreover, the input to the child is not as impoverished as previously thought (e.g., Chomsky, 1965), but is rich with useful information. For instance, infants' constrained visual world subsequently constrains their encoding of linguistic input (e.g., Yu, Smith, Klein, & Shiffrin, 2007). Additionally, regularities within the linguistic environment provide reliable cues for visual object categorization (Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002). Computational modeling approaches have also largely outstripped nativist accounts, as they have construed learning as a stochastic process over distributed representations (e.g., Elman, 1990), as opposed to a manipulation of discrete symbols (e.g., Marcus et al, 1999). More recent claims (e.g., Spencer, Blumberg, McMurray, Robinson, Samuelson, & Tomblin, 2009) have argued that nativist theories should be abandoned entirely, as they ignore epigenetics: The view that developmental processes including language acquisition emerge via cascading interactions across multiple levels of causation, from genes and environment. Also, close examinations of population-level data have shown no evidence for a discontinuity in second language acquisition at the critical period (Hakuta, Bialystok, & Wiley, 2003), which runs in stark contrast to hypotheses regarding a biologically-defined critical period. Given the fragility of nativist accounts, there is room to explore how the language learner harnesses the information available to them in the environment in order to acquire a language.

1.2 The Emergence of Statistical Learning Accounts

Despite the early influence of innate or modularist accounts, experience-dependent theories for language acquisition have gained prominence in recent decades. In one of the most influential findings on the study of language acquisition, Saffran, Newport, and Aslin (1996), and Saffran, Aslin, and Newport (1996) empirically demonstrated that distributional cues inherent to linguistic input may play an important role in segmenting words from fluent speech. The experiments by Saffran and her colleagues were a particularly intriguing endeavor, as one of the first tasks facing the infant language learner is discovering the words embedded in fluent speech. Moreover, word segmentation in a natural language is a complicated task as there are multiple cues that may be confounded with word boundaries, including prosodic cues (Christophe, Dupoux, Bertoncini, & Mehler, 1994; Cutler, 1994; Cutler & Norris, 1988), stress patterns (Echols, 1993; Echols & Newport, 1992, Jusczyk, Houston, & Newsome, 1999), and speakers' tendency to rarely pause between words (Cole & Jakamik, 1980). However, linguists had determined that words could reliably be segmented based on the conditional or transitional probabilities between adjacent syllables (Harris, 1955). Specifically, the transitional probabilities of adjacent syllables within words is higher than the transitional probabilities between words. For instance, in the phrase "*personal computer*", the probability of the syllable "*al*" following "*son*" is higher than the probability of "*com*" following "*al"*. It is this distinction between within and between word transitional probabilities that allows the language learner to discover the boundaries between words, and segment words from fluent speech (e.g., Hayes & Clarke, 1970).

To demonstrate word segmentation using an artificial language, adults (Saffran, Newport, et al., 1996) and 8-month-old infants (Saffran, Aslin, et al., 1996) were exposed to a structured, unsegmented artificial language where the only cue to word boundaries were the transitional probabilities between syllables. Both studies found that even after a brief exposure, participants could readily identify words from the speech stream, and demonstrated that experience with the distributional cues within a language may contribute to language acquisition. However, these studies did not definitively show that transitional probabilities drive word segmentation. To overcome this, Aslin, Saffran, and Newport (1998) constructed a speech stream that did not confound transitional probabilities and joint probabilities. To accomplish this, both the trained words part-word foils used in the test phase were equated in frequency of co-occurrence during the familiarization phase. However, the trained words had higher conditional probability of co-occurrence, or higher transitional probabilities, than the part-word foils. They found that infants could reliably discriminate trained words and part-words, and demonstrated that the computation of transitional probabilities supersedes the simple frequency of co-occurrence between units. This discovery that language learners could detect the distributional regularities within speech was termed statistical language learning, and has since become an expanding area of research on the experience-dependent mechanisms of language acquisition.

The findings from these initial studies on statistical word segmentations coincided with other emergentist theories at the time. Emergentist theories were concerned with how internal representations developed or emerged gradually due to exposure, and included how information was computed across distributed representations based on the principles of associative learning (*see* Bates & Elman, 1996; Cleermans, 1993; Hollich, Hirsh-Pasek, & Golinkoff, 2000). Major developments in infant research demonstrated how experience shaped linguistic perception, including the narrowing of phonetic perception for native-language consonants (Werker & Tees, 1984) and vowels (Kuhl et al., 1992), stress patterns (e.g., Jusczyk, Cutler, & Redanz, 1993), and phonotactics (Jusczyk et al., 1993). These results suggest that infants are skilled learners, and that their perceptual system is organized based as a function of their linguistic experience (*see* Kuhl, 2004; Werker & Curtin, 2005). And, this experience-based learning may not be restricted to infancy: Experience may be capable of guiding statistical learning throughout the lifetime (e.g., Saffran, Newport, & Aslin, 1996; Saffran et al., 1997). There are many other examples of language learners using linguistic input to extract more complex statistical structures, and I will now review some of relevant findings of statistical learning beyond word segmentation.

1.3 Statistical Learning beyond Word Segmentation

1.3.1 Phonetic perception

Different acoustic contrasts are meaningful across different languages. For example, in English, the distinction between /r/ and /l/ phonemes is meaningful, as in “rake” and “lake”, while in Japanese, these phonemes are treated as part of the same category and are thus indistinguishable to Japanese speakers. Learning these distinctions is the result of statistical learning of distributional phonemic information within one’s native language. In a series of studies, Maye and her colleagues (Maye, Werker, & Gerken, 2002; *also see* Yoshia, Pons, Maye, & Werker, 2010) demonstrated that infants or adults form phonemic categories based on experience with distributional regularities. In Maye, Werker, and Gerken (2002), infants were familiarized with sounds along a phonetic /da/-/ta/ continuum that was distributed as either unimodal or bimodal. Specifically, in the bimodal condition, infants heard the endpoints of the /da/-/ta/ continuum more frequently than stimuli at the mid-point of the continuum. In the unimodal condition, on the other hand, infants heard stimuli from the mid-point of the continuum more frequently than stimuli at the endpoints of the continuum. They found that during the test phase, only infants in the bimodal training condition could distinguish between phonetic tokens at the endpoints of the continuum, while those trained in the unimodal condition could not. This finding was significant as it demonstrated that infants are sensitive to the statistical distribution of speech sounds in language input, and that this sensitivity shapes their speech perception. This finding was also consistent with other accounts showing that by 6 months of age, native-language experience has shaped infants' phonetic perception (e.g., Kuhl et al., 1992).

1.3.2 Syntax acquisition

One of the key criticisms from the original Saffran et al. (1996) findings was that it is grammar learning, and not word learning, that is informative for understanding the mechanisms involved in language learning; According to Pinker (1997), the results from Saffran et al. (1996) were seen as an inconclusive account of language acquisition via a distributional learning mechanism. A response to this criticism came from Gómez (2002), who demonstrated that statistical learning can account for learning of higher-level linguistic structures. In their study, adults and 8-month-old infants were trained on an artificial language with an *AXB* structure: The first element in the sequence, *A*, reliably predicted the final element, *B*, while *X* was not predictive of either *A* or *B*. The key manipulation in this study was that *X* came from a variable set size across conditions, with vocabularies of either 2, 6, 12, or 24 words. Gómez (2002) found that infants and adults only learned the non-adjacent *A-B* relationship when *X* was least predictive and came from the largest set size. The findings from Gómez (2002) were particularly influential for statistical learning approaches, as it showed that language learners could use distributional regularities beyond the single-order adjacent relationships originally demonstrated in Saffran, Aslin, and Newport (1996). Further, the demonstration of statistical learning for non-adjacent sequences demonstrated that statistical learning may contribute to learning grammatical or syntactic structures, and was closely related to other implicit learning approaches examining language acquisition, namely, artificial grammar learning paradigms (e.g., Knowlton, Ramus, & Squire, 1992; Knowlton & Squire, 1996;

Matthews et al., 1989; Reber, 1967; 1989, *see* Christiansen, 2018 *and* Perruchet & Pacton, 2006 *for discussions*).

1.3.3 Bridging between linguistic levels

Distributional learning can also be seen when bridging between linguistic levels. In a novel approach, Graf Estes, Evans, Alibali, and Saffran (2007) used a speech segmentation paradigm to train 8-month-old infants on an artificial language. Similar to the Saffran, Newport, and Aslin (1996) paradigm, infants heard a structured, unsegmented speech stream containing four, bi-syllabic, nonsense words. Following language exposure, the trained "words" from the artificial language were used as labels for novel objects. They found that infants who heard labels that were either inconsistent with the regularities in the training language (Experiment 1), or familiar sequences with low internal probabilities (Experiment 2) did not learn the labels. However, infants who heard labels that were consistent with the statistical regularities in the training stimuli learned the labels for the novel objects. This finding was the first to demonstrate that words with a high internal probability in a speech stream may be treated as proto-lexical traces or candidate words, and expanded our understanding of the output of statistical learning. Additionally, it provided convincing support that statistical cues interact (*also see* Lany & Saffran, 2010; Thiessen & Saffran, 2007; Sahni, Seidenberg & Saffran, 2010), with learning at one level affecting learning downstream.

1.3.4 Summary

Despite early criticism to statistical learning approaches (*see* Pesetsky, Wexler, & Fromkin, 1997) it is clear that there are applications for distributional learning beyond the original findings from Saffran, Aslin, and Newport (1996). The findings reviewed here show that statistical learning is valuable for the naïve language learner for learning multiple levels of linguistic structure. In this thesis, I will focus specifically on the statistical word segmentation paradigm originally explored by Saffran, Newport, and Aslin (1996). Crucially, I will focus on the *process* of statistical learning, as much of the prior research has primarily focused on the *outcome* of statistical learning (*see* Romberg & Saffran, 2013). The cognitive processes involved in statistical learning, as well as the constraints on learning, will be examined in detail. In this chapter, I will first explore some of the theoretical background on statistical language learning. Next, there will be a discussion of some of the domain-general and domain-specific constraints on statistical learning, and how statistical learning is situated within our broader cognitive architecture. Finally, some current issues in statistical learning paradigms will be addressed before proceeding to a review of the studies included in this thesis.

1.4 Computational Approaches to Statistical Learning

1.4.1 The critical period and the Less is More Hypothesis

It is important to review the relevant theoretical and computational work that has been influential for the field of statistical learning, as much of this work guides the discussion of the cognitive processes theorized to underlie this process. An early and influential theory in this realm was the Less is More Hypothesis from Newport (1990). The basis for this theory was earlier work from Lenneberg (1967), hypothesizing a biologically-defined critical period for language acquisition. Since Lenneberg's (1967) proposal, however, it was difficult to empirically test for the critical period experimentally given strong (and ethically conscionable) constraints on how scientists can deprive children of input (*cf.* 'natural experiments' involving children raised under extreme cases of sensory deprivation, e.g., Curtiss, 1977). In a ground-breaking finding at the time, Johnson and Newport (1989) examined English second-language proficiency of

Korean children who moved to the United States. They reported that children who moved before the age of 7 achieved better second language proficiency than those who arrived after the age of 7. They then concluded that there is, in fact, a critical period of language acquisition, and that this critical period extended to second language acquisition. This finding also raised an intriguing question: Why were younger children, who have lower cognitive ability, more proficient than older children at learning a second language? To reconcile this, the Less is More Hypothesis (Newport, 1990) suggested that the restricted working memory of children was advantageous for language acquisition, and attempted to offer a causal explanation for this phenomenon. It was thought that the smaller perceptual window in children, due to a smaller working memory capacity, constrained the number of possible analyses language learners can make relative to adults, who have a larger perceptual window. Hence, the likelihood of making a correct analysis is increased as extraneous or complex information is omitted.

1.4.2 Testing the Less is More Hypothesis

Following Newport's (1990) proposal, Elman (1993) tested the Less is More Hypothesis with a model of syntactic agreement acquisition using a simple recurrent network (SRN; Elman, 1990). An SRN is a three-layer feed-forward artificial neural network consisting of input units, hidden units, and output units. After the first item is fed into the network, the state of the hidden units is "copied back" to the context or input units. Activation in the network is propagated forward, and the output is compared to the target via backpropagation. The development of an SRN for natural language processing was a substantial advancement for cognitive and linguistic theory, as it was the first simulation to break away from *a priori* commitments to specific linguistic units or representations (i.e., units for phonemes and words), and instead proposed that units may be an emergent consequence of learning. To test the Less is More Hypothesis, Elman (1993) trained an SRN on a corpora of English sentences. The context units of the model represented working memory, and the size of these units was varied in order to simulate the modulation of working memory capacity throughout development. Elman (1993) found that the network could successfully learn the dependencies within the corpora when a gradual increase in working memory was instantiated in the model, which mirrored developmental increases in working memory. However, when the working memory parameter was initially set at a full capacity, the model failed to learn. This finding was important on two accounts. First, it offered a plausible explanation as to why there are developmental constraints on language acquisition. Second, this model provided a role for working memory within language acquisition.

1.4.3 Memory-based accounts of statistical language learning

One of the most important contributions from Elman's (1993) model was the inclusion of the role of memory in segmentation, and the role for memory in segmentation was clearly acknowledged in later chunking models. Chunking models account for segmentation by grouping speech segments or syllables together into manageable units. One illustrative example of a chunking model adapted for speech segmentation was the PARSER model (Perruchet & Vinter, 1998), which was used to test the behavioural findings reported by Saffran et al. (1996). This model, and later chunking models (e.g., Frank, Goldwater, Griffiths, & Tenenbaum, 2010; Giroux & Rey, 2009; Orbán, Fiser, Aslin, & Lengyel, 2008), relied on three interrelated memory processes for word segmentation: Activation, decay, and interference. When first exposed to a string of syllables, the model randomly groups them into chunks. However, as the model proceeds through the input stream, the activation of chunks stored within memory decrease over time, comprising the decay function. This gradual decay of chunks stored in memory continues, unless the chunks are re-encountered. In this way, groups that are likely to co-occur (e.g., syllables that co-occur within a word) are likely to be chunked and stored as a memory trace due to re-activation. Groups that are less likely to occur (e.g., syllables that span a word boundary) are less likely to be chunked and are thus subject to decay. Similarly, there may be interference for the chunks stored in memory when conflicting segmentations are made. Speech segmentation, then, consists of synthesizing a set of elements (e.g., syllables) into larger units (e.g., words). For instance, in learning the phonological form of the word "*computer*", learners would differ in their response to the subcomponents of the unit, such as "*comp*", "*u*", or "*ter*", as they become more familiar with the overall unit. As learners become more familiar with (or learn) the overall unit, the subcomponents embedded within that unit become less plausible as discrete units. Thus, "*comp*", "*u*", and "*ter*" become less likely as candidate words, as "*computer*" becomes more likely.

The iMINERVA model (Thiessen & Pavlik, 2013) was a recent update of these earlier chunking models, and continued to invoke memory processes in statistical learning. In this memory-based approach, a probe to memory activates prior exemplars based on a function of their similarly to one another: More similar exemplars are activated more strongly than less similar exemplars. If no similar prior exemplar exists, the probe is integrated into memory as a new exemplar. Features that are consistent across exemplars become strengthened over accumulated exposure, while features that are inconsistent are weakened. Thus, the model eventually comes to represent a set of exemplars that are prototypical in nature, or, reflect the central tendency of the distributions to which it has been exposed. Models such as iMINERVA were an important development, as they extended beyond statistical learning of word boundaries to incorporate other types of statistical learning (Thiessen, 2017).

1.4.4 Summary

Treating statistical learning as a memory-based process, as described by these computational approaches, has relevant theoretical implications. First, it provides a link between statistical learning and other forms of learning. For instance, factors that influence memory have also been shown to influence statistical learning, including spaced practice and consolidation (Gómez, Bootzin, & Nadel, 2006), and memory constraints in how stimuli are encoded, stored, and accessed may similarly constrain language learning (Christiansen & Chater, 2015). Returning to the idea that language learning capabilities decrease with development (e.g., Elman, 1993; Newport, 1990), the connection between statistical learning and memory may also provide insights into this developmental change. As discussed by Thiessen (2017), developmental changes in language learning may be characterized by changes in memory abilities, with earlier stages of learning being slow and flexible learning, to later stages where learning is fast but constrained. The models described in this section are by no means an exhaustive description of the computational approaches in statistical learning research. However, they all highlight a role for memory in statistical learning. In the next section, I will explore how memory and other domain-general cognitive processes may be involved in statistical language learning, and how these are contrasted with domain-specific constraints on statistical learning.

1.5 Domain-Generality and Domain-Specificity in Statistical Learning

1.5.1 Terminology

Before proceeding to a review of different domain-general processes that may be involved in statistical learning, it is important to comment on the different conceptualizations of cognitive and linguistic architecture within “domain-specific” and “domain-general” systems. Domain-specificity often refers to Fodorian modules, that is, modules that operate on only certain kinds of input, and that are highly specialized (Fodor, 1983). For the purposes of nativist or modularist approaches, language is viewed as domain-specific, and a distinct and encapsulated cognitive system (e.g., Chomsky, 1965; Levy & Kavé, 1999). Domain-general processes are quite distinct from these encapsulated views, and instead view mental activity as being distributed in nature (e.g., Rumelhart, McClelland, & The PDP Research Group, 1986). In terms of language acquisition accounts, domain-generality was thought to arise from a unitary learning mechanism operating across stimulus types (e.g., Skinner, 1957). However, this account seemed insufficient (*see* Saffran & Thiessen, 2007), and domain-generality in statistical learning has more recently been viewed as a result of domain-general constraints in cognition and perception. Specifically pertaining to statistical learning account, domain-generality in has been ascribed to statistical learning, as similar learning has been shown across stimulus types (*see* Thiessen, 2011). However, domain-generality in statistical learning is not the result of a unitary learning mechanism (e.g., Endress & Mehler, 2009), and there are qualitative and quantitative differences in statistical learning outcomes across domains (Conway & Christiansen, 2005). What is more likely is that statistical learning is comprised of multiple distributed domain-general processes, including memory and attention (Arciuli, 2017), that are responsible for engaging in domain-general computations across stimulus types (Frost, Armstrong, Siegelman, & Christiansen, 2015). Views on domain-specificity have also been updated from modularist accounts, with domain-specificity arising from domain-specific processing constraints across different input types. I will now review some of the evidence for domain-specificity in statistical learning, and possible component cognitive processes that may be involved in statistical learning in a domain-general way.

1.5.2 Domain specificity in perceptual processing

Recent theoretical accounts (e.g., Frost et al., 2015) suggest that domain-specificity in statistical learning arises from specific perceptual processing constraints that distinguish learning across domains. In this view, statistical learning is thought to be subject to constraints that are specific to the modality of input. For instance, auditory information unfolds over time, and is thus subject to the temporal characteristics of the input. Visual information, on the other hand, is not sensitive to the same temporal constraints, as visual information is processed (more or less) instantaneously. Evidence for these domain-specific constraints in statistical learning comes from reports comparing statistical learning across domains that consistently reveal domain-specificity. For instance, rate of presentation has different effects on statistical learning of audio versus visual material (Conway & Christiansen, 2009). Additionally, performance across auditory and visual statistical learning tasks has been shown to be uncorrelated (Siegelman, Bogaerts & Frost, 2017). Further evidence for domain-specificity in statistical learning was clearly demonstrated in Conway and Christiansen (2005). In this study, statistical learning across auditory, visual, and tactile sequences was compared. The researchers found quantitative differences in learning across domains, with better overall memory for auditory sequences, and qualitative differences in the memory for learned sequences across domains. Conway and Christiansen thus demonstrated that statistical learning is constrained by the modality of input.

The domain-specific constraints proposed by Frost et al. (2015) are also supported by the neurobiological distinctions in statistical learning across domains. For example, auditory statistical learning results in increased activation in primary auditory processing regions of the cortex (e.g., Cunillera et al., 2009), while visual statistical learning results in greater activation in primary visual areas (e.g., Turk-Browne, Scholl, Chun, & Johnson, 2009). These findings of domain-specific cortical activation further clarify statistical learning operates in a domain-specific way. However, there are other domain-general cognitive constraints on statistical learning, which I will now explore.

1.5.3 Attention

Although it was originally proposed that statistical learning can operate in the absence of attention (Saffran et al., 1997), subsequent research has provided contrasting evidence to this claim. In an investigation of attentional interference on statistical learning, Toro, Sinnett, and Soto-Faraco(2005) had participants complete a word segmentation task while concurrently engaged in an attention-demanding task that was verbal or nonverbal in nature. The attentional interference task was either auditory (noises) or visual (pictures), or involved attending to pitch changes in the speech stream itself. They found that across all secondary task conditions, diverted attention led to an impairment in statistical word segmentation, leading to the conclusion that statistical learning is reliant on attentional resources. Other work has also shown that when attention is divided (Fernandes, Kolinsky, & Ventura, 2010) or directed to an irrelevant stream (Emberson, Conway, & Christiansen, 2011), statistical learning is disrupted. Attentional interference effects have also been found on visual statistical learning tasks (Turk-Browne, Junge, & Scholl, 2005). It may also be that when attention is enhanced, statistical language learning performance improves. To investigate this, Thiessen, Hill and Saffran (2005) examined infant-directed speech facilitated word segmentation in 7-month-old infants. Prior research suggests that infant-directed speech holds infants' attention better than adult-directed speech (e.g., Werker, Pegg, & McLeod, 1994). Thiessen et al. (2005) found that compared to infants who heard adult-directed speech, infants who heard infant-directed speech were more accurate at identifying words from an artificial language following language exposure, and concluded that the infant-directed speech promoted attention to the statistical regularities in the language. Thus, even though statistical learning occurs implicitly (e.g., Reber, 1967), some degree of attention is may support learning.

1.5.4 Working memory

Some accounts view working memory as a domain-specific system, with separate stores for visuospatial and verbal information, mediated by a domain-general central executive (e.g., Baddeley & Hitch, 1974) and episodic buffer (Baddeley, 2000), while other models view working memory as a domain-general system (Daneman & Carpenter, 1980; 1983; Just & Carpenter, 1992 *also see* Cowan, 1999; 2001). Recently, it has been suggested that domain-general processes within working memory are involved in statistical learning. Arciuli and Simpson (2011) speculated that for a task like statistical learning, where implicit computations are made between sequentially delivered input, it is likely that some form of working memory is involved. This view was later echoed by Janacsek and Nemeth (2015), who also speculated that some form of short-term storage is necessary for extracting statistical structure. Similar to the findings of attentional interference in statistical learning, recent research has shown that working memory interference also disrupts statistical learning. Speech segmentation has been shown to be disrupted via an articulatory suppression task (Lopez-Barroso, de Deigo-Balaguer, Cunillera, Camara, Münte, & Rodríguez-Fornells, 2011), showing that phonological rehearsal within working memory may be important for speech segmentation. Subsequent research has shown that statistical word segmentation is improved with a slower compared to a faster speech rate, which is thought to facilitate maintenance within verbal working memory (Palmer & Mattys, 2016, Experiment 1 and 2). Statistical learning was also shown to be disrupted when participants were engaged in a concurrent verbal or visual working memory task. Additionally, the level of working memory disruption was largest when the artificial language was presented at a slow rate of articulation, or when the contributions from working memory were thought to be greatest (Palmer & Mattys, 2016, Experiment 3). The interference effects from both verbal and visual secondary tasks led Palmer and Mattys (2016) to conclude that statistical learning is supported by domain-general processes within working memory. One of the goals of this thesis is to examine the effects of domain-specific and domain-general working memory interference in a statistical word segmentation task (Chapter 2).

1.5.5 Executive functions

Although there is little research on the subject, some have speculated that executive functions may also be involved in statistical or implicit learning. Executive functions are domain-general processes, and researchers have identified three related but separable components of executive functions, including inhibitory control, shifting, and updating or attentional monitoring (Miyake, Friedman, Emerson, Witzki, Howerter, & Wager, 2000). Kapa and Colombo (2014) examined the relationship between executive functions and artificial grammar learning in adults and preschoolers, and found that the adults' artificial grammar learning was predicted by their inhibitory control abilities (Study 1), and that children's performance was predicted by their attentional monitoring and shifting abilities (Study 2). Importantly, executive functions uniquely predicted performance on the artificial grammar learning tasks after controlling for vocabulary and working memory abilities. This finding is consistent with other implicit learning research showing a positive association between executive functions and performance on procedural learning tasks (e.g., Beaunieux et al., 2006). Converging evidence comes from neuroimaging research demonstrating the involvement in statistical learning of the prefrontal cortex (Robertson, Tormos, Maeda, & Pascual-Leone, 2001), an area considered critical to executive functions (*see* Kane & Engle, 2002). However, further examination of the involvement of executive functions in statistical language learning are necessary before any strong conclusions can be made.

1.5.6 Summary

Taken together, it appears that statistical learning involves multiple processes that are domain-general in nature, including attention, working memory, executive functions, and general memory processes. The notion that statistical learning is comprised of multiple component processes is consistent with a recent account from Arciuli (2017). In her account, the components involved in statistical learning relate to the encoding, integration, and abstraction processes necessary for segmentation (e.g., Thiessen & Pavlik, 2013), and may include attention, processing speed, and working memory. Individuals vary in terms of these underlying components, and statistical learning across different domains may vary as a function of how a specific statistical learning task draws on these underlying components. Thus, if statistical learning is viewed as a multicomponent ability, restrictions or impairments in any one of these processes may constrain statistical learning. More work is needed to examine how variations in domain-general cognitive processes are involved in statistical learning. One of the goals of this thesis is to examine this in terms of domain-general working memory restrictions in statistical word segmentation (Chapter 2). Likewise, I also examine how individual variability in language, working memory, and other cognitive abilities is related to statistical learning on verbal and visual segmentation tasks (Chapter 3).

1.6 Neuroimaging Support for Statistical Learning

Our understanding of the processes involved in statistical learning can be informed by examining its neural bases. There are a few recent functional magnetic resonance imaging (fMRI) studies that have examined the neural bases of auditory statistical learning. Two main regions that have been identified are the superior temporal gyrus (STG), and inferior frontal gyrus (IFG). The STG is involved in auditory perception and speech processing (Hickock & Poeppel, 2004), and has been shown to have increased activation for structured auditory sequences relative to random auditory sequences (McNealy, Mazziotta, & Dapretto, 2006), and has been implicated in the segmentation of words from fluent speech (e.g., Cunillera et al., 2009). The IFG seems to track the expression of auditory statistical learning in behaviour (Karuza, Newport, Aslin, Starling, Tivarus, & Bavelier, 2013). In a natural language learning experiment, Plante, Patterson, Dailey, Kyle, and Fridriksson (2014) examined functional changes in neural activation with accumulated language exposure. They found that when exposed to structured input, adults' brain activation reflected increased activation with a widely distributed language network including the superior gyrus, the temporo-parietal-occipital junction and frontal regions, and that activation levels within this network were related to offline behavioural measures of word segmentation. Overall, structured auditory input leads to reliable and measureable neural responses reflective of language processing.

In addition to the relevant evidence of cortical involvement in auditory statistical learning, the distinction between domain-specific and domain-general constraints on statistical learning is also well-supported by neuroimaging research. Statistical learning invokes stimulus-specific cortical activation, with activation in cortical regions being in accordance with the mode of learning. As mentioned, for auditory statistical learning, activation can be seen in the STG and IFG (e.g., Cunillera et al., 2009; Hickok, 2012; Karuza et al., 2013; Overath et al., 2007), while for visual statistical learning, activation is seen primarily in visual areas including the lateral occipital cortex (e.g., Turk-Browne et al., 2009). The domain-general computations involved in statistical learning may be mediated by memory systems within the medial temporal lobe, specifically, the hippocampus. Evidence for domain-general involvement of the hippocampus in statistical learning comes from a case study of a patient with bilateral hippocampus loss who exhibited deficits in statistical learning across a range of auditory and visual tasks (Schapiro et al., 2014). Given the limitations of a single-case design, however, this possibility would need to be examined further. The hippocampus also has a unique role in statistical learning in that different sub-regions of the hippocampus seem to have distinct roles in extracting distributional regularities. These sub-regions include the dentate gyrus and CA3, which are responsible for encoding of specific exemplars and pattern extraction, respectively (Schapiro et al., 2012). The complementary involvement of these hippocampal sub-regions in statistical learning is consistent with memory-based accounts, which proposed distinctions between the extraction and integration processes involved in statistical learning (e.g., Thiessen & Pavlik, 2013).

Other relevant neuroimaging support comes from studies using electroencephalography (EEG) to examine event-related potentials (ERPs) in response to artificial language stimuli. Different studies have proposed several ERP components that may reflect word segmentation, including the N100 (Sanders, Newport, & Neville, 2002), the P200 (Cunillera, Toro, & Sebastián-Gallés 2006; De Diego-Balaguer, Toro, Rodríguez-Fornells, & Bachoud-Lévi, 2007), P300 (Batterink, Reber, Neville, & Paller, 2015), the N400 (de Diego-Balaguer et al., 2007; Cunillera, Toro, & Rodríguez-Fornells, 2006; Cunillera et al., 2009, Sanders et al., 2002), and the LPC (Batterink et al., 2015). In the first study to examine the ERP components related to word segmentation, Sanders and colleagues (2002) examined the ERP responses to nonsense words in a statistical word segmentation task. Following the artificial language exposure, participants showed an enhanced N100 and N400 in response to the nonsense words. The N400 has commonly been associated with processing meaningful words (e.g., Kutas & Federmier, 2011), which lead Sanders et al., (2002) to suggest that the N400 response following language exposure was reflective of proto-lexical trace formation of the novel words. Subsequent research from Abla, Katahira, and Okanoya (2006) and Cunillera et al. (2009) examined the dynamic nature ERP responses during auditory statistical learning tasks, and found that the amplitude of the N400 response increased with accumulated language exposure. In examining how statistical regularities modulated ERP responses to trained words, Batterink et al. (2015) measured the amplitude of the P300 in response to word onset, medial, or final syllables following exposure to an artificial language. They found that the P300 amplitude linearly increased as a function of syllable position, suggesting that neural responses to trained words differ as a function of learned statistical regularities. What has not been explored is how this modulation in ERP responses to predictable and unpredictable syllables changes as a function of exposure to structured input, which will be examined more fully in this thesis (Chapter 3).

1.7 Relevant Issues in Statistical Learning Research

1.7.1 Modelling language acquisition in infancy with laboratory experiments

Critical periods are biologically determined and fixed time periods in which an organism's neural functioning (and related behaviour) is open to external input and modulation. In the original conception, a critical period was seen as time-invariant, although more recent formulations suggest that some biological developments may be better characterized by a "sensitive period" (e.g., Bateson, 1979; Michel & Moore, 1995). As mentioned previously, the idea of a critical period for language acquisition was first proposed by Lenneberg (1967), who suggested that the process of language acquisition is deeply constrained by biology. Lenneberg's (1967) proposal was seen to be consistent with Chomsky's (1965) argument for a specialized, language-acquisition device. These ideas converged on their biological restrictions for when language can be acquired. Although theories of language acquisition have largely moved away from strict nativist theories, there nevertheless seem to be some aspects of language for which humans seem to be biologically endowed (*see* Werker & Tees, 2005 *for a review*), and age of acquisition maintains a strong predictive relationship with later language proficiency (*see* Newport, Bavelier, & Neville, 2001). Given that there are some aspects of language that seem to be better acquired earlier in life, it raises the question as to why much of the research on statistical learning has been conducted on adults, and whether adults are an appropriate model in which to examine this language acquisition mechanism.

Early findings supported no difference between children and adults on a statistical word segmentation task (Saffran, Newport, Aslin, Tunick, & Barrueco, 1997). However, one caveat distinguishing statistical learning paradigms in infants and adults is that infants are generally trained on simpler artificial languages than adults, with vocabularies of four (e.g., Saffran, Aslin & Newport, 1996) compared to six words (e.g., Saffran, Newport, & Aslin, 1996). Additionally, although earlier reports suggested age invariance, later reports (Arciuli & Simpson, 2011) showed that statistical learning *improves* with age. Although it can be noted that Arciuli and Simpson (2011) used a visual statistical learning task, and this finding cannot be used as direct evidence against language acquisition accounts, it nevertheless raises the possibility that statistical learning in general improves with age. Another possibility is that there is a bidirectional relationship between statistical learning and language acquisition, such that improvements in statistical learning abilities with age boost language learning, and language that is already acquired shapes statistical learning (e.g., Siegelman, Bogaerts, Elazar, Arciuli, & Frost, 2018). However, statistical or implicit learning is underpinned by phylogenetically older brain structures that mature early and remain invariant throughout much of development (Reber, 1992). In line with this, it may be that at least some aspects of statistical learning are intact from infancy to adulthood.

Although examining the developmental trajectory of statistical learning is not examined in this thesis, it is nevertheless a stipulation in the research that is important to address. It seems that there are some statistical learning advantages for older children (e.g., Arciuli & Simpson, 2011), and this may extend into adulthood. Thus, the findings from adult statistical learning paradigms may constrict or confound our interpretations of language acquisition in infancy. However, statistical language learning may remain unchanged through development (e.g., Saffran et al., 1997). Further still, there is a plethora of research examining statistical learning in infancy (*see* Gómez & Gerken, 2000; Romberg & Saffran, 2010 *for a review*) that has already been informative for theoretical accounts of statistical learning. It is therefore likely that investigating statistical learning outside of the critical period is, at minimum, informative for existing theories of language acquisition based on statistical learning research in infants.

1.7.2 Measurement approaches

Typically, in adult statistical language learning paradigms, participants are exposed to a structured, unsegmented speech stream for 21 minutes, and complete a two-alternative forced-choice task following language exposure. This task involves making a forced-choice response between two exemplars, which is presumed to test participants' knowledge of the distributional structure of the language. It has recently been demonstrated that this test paradigm relies on explicit knowledge of the artificial language, and it in fact underestimates implicit learning in statistical learning paradigms (Batterink, Reber, Neville & Paller, 2015). Siegelman, Bogaerts, and Frost (2016) also addressed a number of psychometric limitations in the two-alternative forced-choice test, including there being a limited number of items included at test, most of the sample performing at-chance, and all items being at the same level of difficulty. Earlier work (*see* Romberg & Saffran, 2013) suggested that implicit online measures may more accurately measure the process of statistical learning, and proposed a move away from measuring learning via explicit, post-training outcome measures. Some recent methodologies have been proposed, including using reaction time measures in response to trained language stimuli (e.g., Batterink et al., 2015; Batterink, Reber, & Paller, 2015), measuring event-related potentials (ERPs) during a statistical learning task (e.g., Abla et al., 2006; Cunillera et al., 2006; Cunillera et al., 2009), and measuring neural entrainment during a word segmentation task (Batterink & Paller, 2017). The development of these new measurement approaches is greatly informative for further uncovering the processes involved in statistical learning.

1.8 Objectives and Overview

The central objective of this thesis is to examine the process of statistical language learning. While most prior studies have primarily focused on the product or outcome of statistical learning, the studies in this thesis will examine both the cognitive and neural processes involved in statistical word segmentation, and how statistical learning unfolds in real-time. By combining behavioural, clinical, and neuroimaging approaches, I consider both the cognitive processes speculated to be involved in statistical learning, and measurement approaches that best capture this learning process.

Chapter 2 examines the effects of both domain-specific and domain-general working memory interference on a statistical word segmentation task. Prior studies have shown that dual-task paradigms can impair statistical learning due to attentional interference (e.g., Toro et al., 2005), or domain-general working memory interference (Palmer & Mattys, 2016). To expand on this prior research, Chapter 2 investigates how visual and verbal working memory tasks interfered with performance on a statistical word segmentation. This paradigm allowed me to examine both domain-general and domain-specific working memory interference effects on a statistical word segmentation task.

Chapter 3 extends this investigation on the domain-general and domain-specific constraints on statistical learning. In this study, children with and without a marked language impairment complete both a statistical language learning and visual statistical learning task, as well as a series of linguistic and domain-general cognitive measures. This experiment expands on the currently scarce area of research examining statistical word segmentation in those with a developmental language impairment. Moreover, this study is novel in that it compared performance across a verbal and visual statistical learning task in a clinically impaired group.

A perennial issue in the statistical learning literature is the apparent disconnect from using explicit tasks, such as a word-recognition task, in order to measure implicit learning outcomes. I also speculate that the findings exhibited in Chapters 2 and 3 may be more nuanced than the outcome measures allowed. In order to overcome this limitation, Chapter 4 investigates the application of an implicit, indirect measure of statistical learning. In this study, participants' event-related potentials (ERPs) were measured online during a statistical word segmentation task. Specifically, I examined responses to predictable syllables during language exposure. In this novel approach, I examined how neural responses were modulated in an online fashion as a function of statistical learning.

The findings presented in this thesis will inform theories regarding the domain-general and domain-specific constraints on statistical language learning, and advance the measurement approaches applied to statistical learning research.

1.9 References

Abla, D., Katahira, K., & Okanoya, K. (2008). On-line assessment of statistical learning by event-related potentials. *Journal of Cognitive Neuroscience*, *20*(6), 952-964.

Arciuli, J. (2017). The multi-component nature of statistical learning. *Philosophical Transactions of the Royal Society of London B*, *372*(1711), 20160058.

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

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*(4), 321-324.

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in Cognitive Sciences*, *4*(11), 417-423.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of Learning and Motivation*, *8*, 47-89.

Bates, E., & Elman, J. (1996). Learning rediscovered. *Science*, *274*(5294), 1849-1850.

Bateson, P. (1979). How do sensitive periods arise and what are they for?. *Animal Behaviour*, *27*, 470-486.

Batterink, L. J., & Paller, K. A. (2017). Online neural monitoring of statistical learning. *Cortex*, *90*, 31-45.

Batterink, L. J., Reber, P. J., & Paller, K. A. (2015). Functional differences between statistical learning with and without explicit training. *Learning & Memory*, *22*(11), 544-556.

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language*, *83*, 62-78.

Beaunieux, H., Hubert, V., Witkowski, T., Pitel, A. L., Rossi, S., Danion, J. M., ... & Eustache, F. (2006). Which processes are involved in cognitive procedural learning?. *Memory*, *14*(5), 521-539.

Chomsky, N. (1959). Verbal behavior.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT press.

Christiansen, M. H. (2018). Implicit statistical learning: a tale of two literatures. *Topics in Cognitive Science*.

Christiansen, M. H., & Chater, N. (2015). The language faculty that wasn't: A usage-based account of natural language recursion. *Frontiers in Psychology*, *6*, 1182.

Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *The Journal of the Acoustical Society of America*, *95*(3), 1570-1580.

Cleeremans, A. (1993). Attention and awareness in sequence learning. In *Proceedings of the 15th Annual Conference of the Cognitive Science Society,* 330-335.

Cole, R. A. and Jakamik, J. (1980) A model for speech perception. In *Perception and Production of Fluent Speech,* R. A. Cole. (Ed.) Hillsdale, NJ: Lawerence Erlbaum Associates, Inc.

Conway, C. M., & 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*(4), 561-580.

Cowan, N. (1999). An embedded-processes model of working memory. *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*, *20*, 506.

Cowan, N. (2001). Metatheory of storage capacity limits. *Behavioral and Brain Sciences*, *24*(1), 154-176.

Cunillera, T., Càmara, E., Toro, J. M., Marco-Pallares, J., Sebastián-Galles, N., Ortiz, H., ... & Rodríguez-Fornells, A. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *Neuroimage*, *48*(3), 541-553.

Cunillera, T., Toro, J. M., Sebastián-Gallés, N., & Rodríguez-Fornells, A. (2006). The effects of stress and statistical cues on continuous speech segmentation: an event-related brain potential study. *Brain Research*, *1123*(1), 168-178.

Curtiss, S. (1988). Abnormal language acquisition and grammar: Evidence for the modularity of language. *Language, Speech, and Mind. London, NY: Routledge*.

Cutler, A. (1994). Segmentation problems, rhythmic solutions. *Lingua*, *92*, 81-104.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(1), 113.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450-466.

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(4), 561.

De Diego-Balaguer, R., Toro, J. M., Rodríguez-Fornells, A., & Bachoud-Lévi, A. C. (2007). Different neurophysiological mechanisms underlying word and rule extraction from speech. *PLoS One*, *2*(11), e1175.

Echols, C. H. (1993). A perceptually-based model of children's earliest productions. *Cognition*, *46*(3), 245-296.

Echols, C. H., & Newport, E. L. (1992). The role of stress and position in determining first words. *Language acquisition*, *2*(3), 189-220.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*(3968), 303-306.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*(2), 179-211.

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, *48*(1), 71-99.

Emberson, L. L., Conway, C. M., & Christiansen, M. H. (2011). Timing is everything: Changes in presentation rate have opposite effects on auditory and visual implicit statistical learning. *The Quarterly Journal of Experimental Psychology*, *64*(5), 1021-1040.

Endress, A. D., & Mehler, J. (2009). Primitive computations in speech processing. *The Quarterly Journal of Experimental Psychology*, *62*(11), 2187-2209.

Fernandes, T., Kolinsky, R., & Ventura, P. (2010). The impact of attention load on the use of statistical information and co-articulation as speech segmentation cues. *Attention, Perception, & Psychophysics*, *72*(6), 1522-1532.

Fodor, J. A. (1983). The Modularity of Mind. MIT Press. Cambridge, Massachusetts

Frank, M. C., Goldwater, S., Griffiths, T. L., & Tenenbaum, J. B. (2010). Modeling human performance in statistical word segmentation. *Cognition*, *117*(2), 107-125.

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

Giroux, I., & Rey, A. (2009). Lexical and sublexical units in speech perception. *Cognitive Science*, *33*(2), 260-272.

Gleitman, L. R., & Wanner, E. (1982). *Language acquisition: The State of the Art*. CUP Archive.

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

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

Gómez, R. L., Bootzin, R. R., & Nadel, L. (2006). Naps promote abstraction in language-learning infants. *Psychological Science*, *17*(8), 670-674.

Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, *18*(3), 254-260.

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological Science*, *14*(1), 31-38.

Harris, Z. S. (1955). From phoneme to morpheme. *Language*, *31,* 190-222.

Hayes, J. R. & Clark, H. H. (1970). Experiments in the segmentation of an artificial speech analog. *Cognition and the Development of Language*. J. R. Hayes (Ed.). New York, NY: Wiley.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, *13*(2), 135.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, *92*(1-2), 67-99.

Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). What Does it Take to Learn a Word?. *Monographs of the Society for Research in Child Development*, *65*(3), 1-16.

Janacsek, K., & Nemeth, D. (2015). The puzzle is complicated: when should working memory be related to implicit sequence learning, and when should it not? (Response to Martini et al.). *Cortex*, *64*, 411-412.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60-99.

Jusczyk, P. W. (2002). How infants adapt speech-processing capacities to native-language structure. *Current Directions in Psychological Science*, *11*(1), 15-18.

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, *64*(3), 675-687.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*(3-4), 159-207.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review*, *99*(1), 122.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, *9*(4), 637-671.

Kapa, L. L., & Colombo, J. (2014). Executive function predicts artificial language learning. *Journal of Memory and Language*, *76*, 237-252.

Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., & Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain and Language*, *127*(1), 46-54.

Knowlton, B. J., & Squire, L. R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(1), 169.

Knowlton, B. J., Ramus, S. J., & Squire, L. R. (1992). Intact artificial grammar learning in amnesia: Dissociation of classification learning and explicit memory for specific instances. *Psychological Science*, *3*(3), 172-179.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences*, *97*(22), 11850-11857.

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience*, *5*(11), 831.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*(5044), 606-608.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621-647.

Lany, J., & Saffran, J. R. (2010). From statistics to meaning: Infants' acquisition of lexical categories. *Psychological Science*, *21*(2), 284-291.

Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice*, *2*(12), 59-67.

Levy, Y., & Kavé, G. (1999). Language breakdown and linguistic theory: A tutorial overview. *Lingua*, *107*(1-2), 95-143.

Lopez-Barroso, D., de Diego-Balaguer, R., Cunillera, T., Camara, E., Münte, T. F., & Rodríguez-Fornells, A. (2011). Language learning under working memory constraints correlates with microstructural differences in the ventral language pathway. *Cerebral Cortex*, *21*(12), 2742-2750.

Marcus, G. F., & Berent, I. (2003). Are there limits to statistical learning?. *Science*, *300* (5616), 53-55.

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

Mathews, R. C., Buss, R. R., Stanley, W. B., Blanchard-Fields, F., Cho, J. R., & Druhan, B. (1989). Role of implicit and explicit processes in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(6), 1083-1100.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), B101-B111.

McNealy, K., Mazziotta, J. C., & Dapretto, M. (2006). Cracking the language code: neural mechanisms underlying speech parsing. *Journal of Neuroscience*, *26*(29), 7629-7639.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*(2), 143-178.
Michel, G. F., & Moore, C. L. (1995). *Developmental Psychobiology: An Interdisciplinary Science*. MIT Press.
Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49-100.
Nazzi, T., & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication*, *41*(1), 233-243.
Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, *14*(1), 11-28.
Newport, E. L., Bavelier, D., & Neville, H. J. (2001). Critical thinking about critical periods: Perspectives on a critical period for language acquisition. *Language, Brain and Cognitive Development: Essays in Honor of Jacques Mehler*, 481-502.
Orbán, G., Fiser, J., Aslin, R. N., & Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences*, *105*(7), 2745-2750.
Overath, T., Cusack, R., Kumar, S., Von Kriegstein, K., Warren, J. D., Grube, M., ... & Griffiths, T. D. (2007). An information theoretic characterisation of auditory encoding. *PLoS biology*, *5*(11), e288.
Palmer, S. D., & Mattys, S. L. (2016). Speech segmentation by statistical learning is supported by domain-general processes within working memory. *The Quarterly Journal of Experimental Psychology*, *69*(12), 2390-2401.
Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233-238.
Perruchet, P., & Vinter, A. (1998). PARSER: A model for word segmentation. *Journal of Memory and Language*, *39*(2), 246-263.
Pesetsky, D., Wexler, K., & Fromkin, V. (1997). Acquiring language. *Science*, *27* (5316), 1177-1181.
Piaget, J. (2005). *Language and Thought of the Child: Selected Works vol 5*. Routledge.
Pinker, S. (1997). Acquiring language. *Science, 27 6*(5316): 1178
Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and brain sciences*, *13*(4), 707-727.
Pinker, S. (1984). *Language learning and language development*. Cambridge, Mass.: Harvard University Press.
Plante, E., Patterson, D., Dailey, N. S., Kyle, R. A., & Fridriksson, J. (2014). Dynamic changes in network activations characterize early learning of a natural language. *Neuropsychologia*, *62*, 77-86.
Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of verbal learning and verbal behavior*, *6*(6), 855-863.
Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of experimental psychology: General*, *118*(3), 219.
Reber, A. S. (1992). The cognitive unconscious: An evolutionary perspective. *Consciousness and cognition*, *1*(2), 93-133.

Robertson, E. M., Tormos, J. M., Maeda, F., & Pascual-Leone, A. (2001). The role of the dorsolateral prefrontal cortex during sequence learning is specific for spatial information. *Cerebral Cortex*, *11*(7), 628-635.

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(6), 906-914.

Romberg, A. R., & Saffran, J. R. (2013). All together now: Concurrent learning of multiple structures in an artificial language. *Cognitive Science*, *37*(7), 1290-1320.

Rumelhart, D. E., McClelland, J. L., & PDP Research Group. (1986). *Parallel Distributed Processing, Experiments in the Microstructure of Cognition.* Cambridge, MA: MIT Press.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926-1928.

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

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological science*, *8*(2), 101-105.

Saffran, J. R., & Thiessen, E. D. (2007). Domain-general learning capacities. *Blackwell Handbook of Language Development*, 68-86.

Sahni, S. D., Seidenberg, M. S., & Saffran, J. R. (2010). Connecting cues: Overlapping regularities support cue discovery in infancy. *Child Development*, *81*(3), 727-736.

Sanders, L. D., Newport, E. L., & Neville, H. J. (2002). Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. *Nature Neuroscience*, *5*(7), 700.

Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M., & Turk-Browne, N. B. (2014). The necessity of the medial temporal lobe for statistical learning. *Journal of Cognitive Neuroscience*, *26*(8), 1736-1747.

Schapiro, A. C., Kustner, L. V., & Turk-Browne, N. B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Current Biology*, *22*(17), 1622-1627.

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

Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., & Frost, R. (2018). Linguistic entrenchment: Prior knowledge impacts statistical learning performance. *Cognition*, *177*, 198-213.

Skinner, B. F. (1957). *Verbal behavior*. BF Skinner Foundation.

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, *13*(1), 13-19.

Spencer, J. P., Blumberg, M. S., McMurray, B., Robinson, S. R., Samuelson, L. K., & Tomblin, J. B. (2009). Short arms and talking eggs: Why we should no longer abide the nativist–empiricist debate. *Child Development Perspectives*, *3*(2), 79-87.

Thiessen, E. D. (2011). Domain general constraints on statistical learning. *Child Development*, *82*(2), 462-470.

Thiessen, E. D. (2017). What's statistical about learning? Insights from modelling statistical learning as a set of memory processes. *Philosophical Transactions of the Royal Society of London B*, *372*(1711), 20160056.
Thiessen, E. D., & Pavlik Jr, P. I. (2013). iMinerva: a mathematical model of distributional statistical learning. *Cognitive Science*, *37*(2), 310-343.
Thiessen, E. D., & Saffran, J. R. (2007). Learning to learn: Infants' acquisition of stress-based strategies for word segmentation. *Language Learning and Development*, *3*(1), 73-100.
Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, *7*(1), 53-71.
Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, *97*(2), B25-B34.
Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*(4), 552.
Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development*, *1*(2), 197-234.
Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*(1), 49-63.
Werker, J. F., & Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, *46*(3), 233-251.
Werker, J. F., Pegg, J. E., & McLeod, P. J. (1994). A cross-language investigation of infant preference for infant-directed communication. *Infant Behavior and Development*, *17*(3), 323-333.
Yoshida, K. A., Pons, F., Maye, J., & Werker, J. F. (2010). Distributional phonetic learning at 10 months of age. *Infancy*, *15*(4), 420-433.
Yu, C., Smith, L. B., Klein, K. A., & Shiffrin, R. M. (2007, January). Hypothesis testing and associative learning in cross-situational word learning: Are they one and the same?. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 29, No. 29).

Chapter 2: Domain-Specific and Domain-General Constraints on Statistical Language Learning

2.1 Introduction

Statistical learning refers to the process of uncovering the regularities within the environment, and has been shown to play a role in learning across a variety of stimulus types. Evidence of statistical learning has been shown in the learning of linguistic (Saffran, Aslin, & Newport, 1996) and non-linguistic (Gebhart, Newport, & Aslin, 2009) auditory sequences, as well as shape (Fiser & Aslin, 2001), spatial location (Mayr, 1996), and tactile sequences (Conway & Christiansen, 2005). Given that statistical learning seems to operate across a diverse array of stimulus types, it has been proposed as a domain-general explanation for human learning (Evans, Saffran, & Robe-Torres 2009; Fiser & Aslin, 2001; Kirkham, Slemmer, & Johnson, 2002; Reber, 1967). This domain-general assumption has been particularly attractive for theories of language acquisition, as it provides evidence against innate, language-specific learning mechanisms (e.g., Chomsky, 1965; Gleitman & Wanner, 1982; Pinker, 1984). Indeed, statistical learning has widely been applied to the study of language learning, including processes related to phoneme discrimination (Maye, Werker, & West, 2002), word segmentation (Saffran, Aslin, and Newport, 1996; Aslin, Saffran & Newport, 1998), word-object pairing (Graf Estes, Evans, Alibali, & Saffran, 2007), and grammar learning (e.g., Dienes, Altmann, Kwan, & Goode, 1995; Gómez, 2002; Gómez & Gerken, 1999). Although there is considerable evidence of statistical learning across a variety of input types, this does not necessarily imply that statistical learning is as unitary mechanism that operates across domains. Critically, the empirical findings examining whether statistical learning is mediated by domain- or modality-specific processes (Conway & Christiansen 2005, 2006; Johansson, 2009), or by domain-general processes such as working-memory (Palmer & Mattys, 2016) or attention (Toro, Sinnett, & Soto-Faraco, 2005) are contradictory. To help reconcile these findings, the present study investigated the domain-specific and domain-general constraints on statistical learning by examining the interference effects of concurrent domain- or language-specific (verbal) or non-language-specific or domain-general (visuospatial) tasks on statistical word segmentation.

In an initial demonstration of statistical language learning, Saffran, Aslin, and Newport (1996) exposed infants to a structured, unsegmented stream of syllables for two minutes. The syllables were concatenated into four nonsense words. Although there were no acoustic or prosodic cues to indicate word boundaries, the syllables within a word always appeared within the same sequence. As a result, the transitional probabilities of adjacent syllables within words (1.0) was higher than the transitional probabilities between words (0.33). Following language exposure, infants showed that they were sensitive to these statistics. In a subsequent study, Aslin, Saffran, and Newport (1998) demonstrated that infants were not only sensitive to not only the simple co-occurrence of frequent pairs, but that they were uniquely sensitive to the transitional or conditional probabilities between syllables. Taken together, these and subsequent findings have provided a demonstration of a mechanism by which language learners can segment words from fluent speech; a fundamental process in language-learning (*see* Romberg and Saffran, 2010).

Considerable research attention has been focused on the mechanisms and constraints involved in statistical learning since Saffran, Aslin, and Newport's (1996) seminal work. Much of this research has begun to examine whether the process of statistical language learning is supported by a language learning subsystem, shaped by language-specific constraints, or by domain-general learning constraints that operates across a variety of input types. Domain-general constraints on statistical language learning would arise from general characteristics of the human cognitive system that shape learning generally, such as working memory (e.g., Newport, 1990), attention (e.g., Thiessen, Hill, & Saffran, 2005), and perception (e.g., Creel, Newport, & Aslin, 2004). Evidence for domain-general constraints on statistical learning has been mounting, particularly from studies examining interference effects on statistical learning across domains. In a recent examination, Palmer and Mattys (2016) had participants complete a word segmentation task while concurrently engaged in a verbal or visual *n*-back task (Experiment 3). They found that both the verbal and visual forms of the secondary task interfered with word segmentation, and inferred that domain-general working memory processes support statistical language learning. Other cross-domain interference effects on statistical learning have demonstrated reduced learning with concurrent auditory attention (Toro et al., 2005), picture-matching (Toro et al., 2005; Fernandes, Kolinsky & Ventura, 2010), and a pitch change detection tasks (Toro et al., 2005). Taken together, these findings provide convincing evidence of constraints on statistical learning arising from general limitations in the human cognitive architecture.

Computational models of statistical learning have attempted to account for the disruption cross-domain tasks have on statistical learning from a capacity-limited framework. In their model of sequential learning, Keele, Ivry, Mayr, Hazeltine, and Heuer, (2003) proposed a multidimensional learning system that is responsible for the learning of complex sequences, such as those computations necessary for word segmentation. Learning through this system occurs automatically, but only attended signals are learned. When processing resources are otherwise engaged, the learning outcomes of the multidimensional system are underspecified. Thus, domain-general interference is possible under this model when there are insufficient processing resources to adequately attend to the to-be-learned stimuli. Similarly, memory based models of statistical learning, such as iMINERVA (Thiessen & Pavlik, 2013), PARSER (Perruchet & Vinter, 1998), TRACX (French, Addyman & Mareschal, 2011), and TRACX2 (Mareschal & French, 2017) describe the mechanisms of statistical learning as being similar to the mechanisms studied in traditional memory paradigms, including how stimuli are encoded, stored, and accessed in memory (*see* Thiessen, 2017). It follows, then, that changes in the precision with which newly extracted information is encoded in memory, particularly through memory interference or decay, changes the nature of the representation itself. As such, modelling efforts have provided a mechanistic stipulation by which domain-general interference in statistical language learning may occur.

Another line of evidence favouring a domain-general view of statistical learning comes from similar findings of learning transitional probabilities across a wide range of stimuli, from speech (Saffran, Aslin, & Newport, 1996), to non-linguistic auditory (Teinonen et al., 2009), visual (Fiser & Aslin, 2001; Turk-Browne, Jungé, & Scholl, 2005), and tactile stimuli (Conway & Christiansen, 2005). However, a reconsideration of this supposition is warranted. First, the findings of similar statistical learning outcomes observed across different stimulus types does not necessarily implicate the involvement of a singular underlying cognitive mechanism. Conway and Christiansen (2005) sought to compare statistical learning across auditory, visual, and tactile modalities in a single study while maintaining comparable experimental procedures across the three paradigms to ensure valid comparisons. Their findings revealed qualitative and quantitative differences in learning across the three modalities, suggesting there are modality-specific constraints on statistical learning. Specifically, auditory statistical learning had a quantitative advantage over visual and tactile statistical learning. In terms of qualitative differences, compared to learning in visual or tactile domains, auditory statistical learning afforded better learning of sequence-final elements. Related work from Seitz and colleagues (2007) examining multi-modal statistical learning found that when participants were presented with a multimodal audiovisual sequence, they could detect distinct sequences from the audio and visual streams. This suggests that the resultant knowledge statistical learning is bound to the modality of exposure, which is consistent with a modality-specific view of statistical learning, and does not result in an amodal representation as would be predicted by a domain-general account (e.g., Altmann, Diens, & Goode, 1995; Peña, Bonatti, Nespor, & Mehler, 2002; Shanks, Johnstone, & Staggs, 1997). Similarly, the knowledge resulting from artificial grammar learning paradigms appears to be stimulus-specific (Conway & Christiansen, 2006; Johansson, 2009). Subsequent neuroimaging evidence has demonstrated that implicit sequence learning seems to activate brain areas dependent on the specific demands of the task (Conway & Pisoni, 2008). This, in turn, raises the possibility that modality-specific neural networks do play a role in statistical learning.

Consistent with these modality specific constraints on statistical learning, previous research has provided evidence for domain-specific interference effects in statistical learning (Noonan, 2014). Although previous studies using interference paradigms have demonstrated domain-general interference effects (e.g., Fernandes et al., 2010; Palmer & Mattys, 2016; Toro et al., 2005), it is possible the secondary tasks impose high cognitive demands, resulting in overarching effects on statistical learning regardless of secondary task domain. Furthermore, the cognitive demand across secondary task domains varied in previous studies, making conclusions about the underlying effects untenable. To reconcile this, Noonan (2014) investigated the interference of a secondary task on statistical learning wherein cognitive load across secondary task domains was consistently varied. Adult participants were exposed to an artificial language similar to that of Saffran, Newport and Aslin (1996) for 28 minutes while they concurrently were engaged in a verbal or visual *n*-back task that was posed either a low (0-back) or high (2-back) demand. It was found that individuals completing a concurrent verbal task, regardless of task load, were impaired on the statistical learning task. Those completing a concurrent visuospatial 0- or 2-back task, however, performed similarly with controls. This provided clear evidence of domain-specific, but not domain-general interference impacted statistical learning, and raised the possibility that statistical learning is supported by domain-specific mechanisms, contrary to previous studies examining interference effects. However, the language exposure in Noonan (2014) was longer than in previous studies of statistical word segmentation, which may have allowed for the learning to be compensated despite secondary task interference. A reduction of the language exposure time, and thus less redundancy in the to-be-learned material, may bring forth a different pattern of interference effects.

In one of the first attempts to reconcile findings related to domain-general and modality-specific constraints on statistical learning, Frost, Armstrong, Siegelman, and Christiansen (2015) suggested that statistical learning is mediated by domain-general neurobiological mechanisms and computational principles, operating across modalities, whilst constrained by the modality of the incoming stimulus. For instance, the encoding of auditory information such as speech unfolds over time, and thus the binding of sequential elements in time is critical in the encoding of incoming auditory information.

Conversely, the visual modality is not bound by the same constraint. These differences are reflected in how the auditory and visual cortices encode to-be-learned material (Chen & Vrooman, 2013; Recanzone, 2009; Salminen, Aho, & Sams, 2013). The proposal that learning is shaped by the modality in which it occurs but is also reliant on an amodal learning mechanism provides possible explanations for both domain-general and modality- or domain-specific interference effects in statistical learning. Specifically, modality-specific interference effects could reflect processing limitations within that domain. Conversely, observed domain-general interference effects could be explained by interference within shared neural networks supporting statistical learning. These may be housed within the medial temporal lobe, specifically, the hippocampus. This region has been shown to be active in a variety of statistical learning tasks across different domains (Schapiro, Gregory, Landau, McCloskey, & Turk-Browne, 2014). Grounded in neuroscience, this framework unifies findings of domain-general and domain- and modality-specific constraints on statistical learning, and provides an adequate description to advance the understanding of the underlying mechanisms that support and constrain statistical learning.

2.1.1 Present study

The involvement of cognitive constraints on statistical learning were examined using an interference paradigm adopted from studies of working memory. Participants were presented with an auditory statistical language learning task while they completed a secondary *n*-back task that was verbal or visuospatial, and imposed either a low- (0-back) or high (2-back) cognitive load. The nature of the secondary *n*-back tasks used in the present study motivated the hypotheses involving the possible domain-general and domain-specific constraints on statistical language learning. The *n*-back task, where $n > 1$, has been described as a "gold standard" working memory task in functional neuroimaging studies (*see* Kane & Engle, 2002). If statistical learning is impaired for participants engaged in a secondary 2-back task across task domains, but not the 0-back tasks, this would point to the involvement of working memory resources in statistical language learning. Such a finding would provide support for descriptions of statistical learning instantiated upon general human cognitive architecture with the involvement of memory-related processes (e.g., French, Addyman, Mareschal, 2011; Keele et al., 2003; Marexchal & French, 2017; Newport, 1990; Palmer & Mattys, 2016; Thiessen & Pavlik, 2013; Toro et al., 2005; Perruchet & Vinter, 1998).

An alternative hypothesis is related to domain-specific interference effects on statistical language learning: Impaired statistical learning performance for participants engaged in a secondary verbal task, but not a secondary visuospatial task, would indicate that statistical learning is reliant on domain-specific language processing resources. This finding would be consistent with previous research (Noonan, 2014) and would support accounts of statistical learning which suggest a reliance on domain- or modality-specific processing constraints (e.g., Conway and Christiansen, 2005, 2006; Johansson, 2009). One consideration when using a visually presented verbal task, as is the case in the present paradigm, is that modality-specific interference effects cannot be assumed. However, considerable evidence from the field of working memory has demonstrated that visual-verbal information is recoded and stored phonologically, resulting in interference with recall of auditorally presented verbal stimuli (Baddeley, 2012). Following from Oberauer and Kliegl's (2006) model of capacity limits in working memory, this interference results from the overlap in the phonological features of the verbal stimuli, regardless of the modality of presentation. It would follow, then, that interference from the visual-verbal (i.e., letter monitoring) task could interfere with statistical language learning if shared phonological resources are encoding both auditory and visual-verbal (written) information, a notion investigated in the present study. Thus, interference effects on statistical language learning from the verbal *n*-back tasks can be presumed to be related to domain-specific capacity limitations.

2.2 Method

2.2.1 Participants

Participants in the present study consisted 104 adults (M_{age} = 18.89 years, SD_{age}= 0.67, N_{male} = 41). Participants received course credit or $10.00 for study participation. All participants reported being monolingual English speakers and had no uncorrected vision or hearing difficulties. Ethics approval for all study procedures and materials was obtained by the University of Western Ontario Non-Medical Research Ethics Board, and written informed consent was obtained from all study participants.

2.2.2 Procedure

Testing took place in a single session within a quiet testing room. The task was administered individually via a laptop computer. The statistical learning task involved a listening phase followed immediately by a test phase. Participants completed the 7-minute listening phase while engaged in one of five concurrent secondary task conditions: (1) Control (no secondary task); (2) verbal, 0-back; (3) verbal, 2-back; (4) visuospatial, 0-back; (5) visuospatial, 2-back. Participants were quasi-randomly assigned to one of the five conditions such that roughly equal numbers completed each condition with no participant factors determining group assignment.

2.2.2.1 Listening phase: Artificial language exposure

Artificial language stimuli. The artificial language employed in the present study was based on the stimuli described by Saffran, Newport, and Aslin (1996). The language consisted of four consonants (p, t, b, d) and three vowels (a, i, u) which, when combined, rendered an inventory of 12 CV syllables. These syllables were then combined to create six trisyllabic “words” in an artificial language: *patubi, tutibu, babupu, bupada, dutaba, pidadi*. Some syllables from the inventory occurred more often within the language than others (e.g., bu occurs in three words, whereas ti occurs in one word). Approximating the transitional probabilities in English, transitional probabilities of syllables varied, and were higher within words (Range: 0.33 to 1.0) than across word boundaries (Range: 0.1 to 0.2) (note these ranges may be even closer in a real language).

Recording the artificial language stimuli. Unlike Saffran, Newport, and Aslin’s (1996) synthesized stimuli, the artificial language in the present study was constructed from audio recordings of a female native-English speaker using a neutral vocal effort. Although much of the available evidence is based on findings from exposure to synthetic speech samples, more recent work of Saffran and colleagues (Graf, Estes, Evans, Alibali, & Saffran, 2007) and similar studies (Lew-Williams, Pelucchi, & Saffran, 2011; Pelucchi, Hay, & Saffran, 2009) have employed naturally produced speech and found similar effects.

Recordings were made in a double walled IAC sound booth with a pedestal microphone (AKG C 4000B) located approximately 30cm from the speaker’s mouth and routed to a USBPre 2 pre-amplifier (Sound Devices) using SpectraPlus software (Pioneer Hill Software, 2008). Recordings were made of each of the 12 target syllables in the middle of a three-syllable sequence, within every co-articulation context required for the language. For example, the syllable *tu* occurred in two words in the artificial language, *tutibu* and *patubi.* For the word *tutibu* in the continuous artificial language stream, the word-initial syllable *tu* could be preceded by the word-final syllables for the remaining five words, *bi, pu, da, ba,* or *di*, and followed only by *ti*. Thus, recordings of these six iterations were made. Alternatively, for the word *patubi* the word-medial syllable *tu* would be preceded only by the word-initial syllable *pa* and the word-final syllable *bi.* Thus, recordings of this one iteration were made. Eight repetitions of each sequence were recorded, and the token with the most neutral pitch contour and best sound quality was chosen and uploaded into Sound Forge Audio Studio (Sony Creative Software) editing software.

Creating the artificial language. Middle syllables from the recorded tokens were extracted by identifying the final offset of vowel oscillation in the previous syllable to the offset of vowel oscillation in the target syllable. The continuous artificial language stream was created by concatenating the medial syllables to create random sequences of the words. In this way, all syllables were spliced together in the same way throughout the entire language regardless of whether the syllables were within a word or across word boundaries. The language maintained a consistent speech rate (average 3.1 syllables/s) using a time stretch and was normalized to a pitch of F0 = 196 Hz using Sound Forge Audio Studio (Sony Creative Software). There were no pauses between words. As such, there were no acoustic cues to word boundaries. The artificial language was comprised of 120 tokens of each of the six words occurring in a random order, with the constraint that the same word never occurred twice in a row.

Listening phase procedure. The listening phase involved exposure to 7 minutes of the artificial language administered via personal headphones. Following Saffran, Newport, and Aslin (1996), participants were told they would hear a nonsense language. No information was provided about the length or the number of words within the language. Those in the control (no secondary task) condition were seated in front of a computer displaying stimuli from the secondary task, but were not instructed to attend to or perform memory operations on the stimuli. Those in one of the four secondary task conditions were instructed to complete the n-back task, and this was highlighted to them as the primary task. This deliberate use of vague instructions regarding the artificial language was done to minimize the chance of participants trying to explicitly learn the language during the experiment.

2.2.2.2 Concurrent secondary tasks

Secondary task stimuli. The n-back task employed in the present study involved presenting one of six alphabetic letters (P, G, T, K, W, C) in 72-point sans-sheriff black font on a white background. Letter case was randomized across trials to avoid reliance on visual recognition of the letter instead of a verbal label when required. The letter on a given trial appeared at one of 6 pseudorandom positions on the screen that were not easily coded verbally (i.e., “top right” or “center” positions were avoided). Letter name and position were counterbalanced across trials so that each letter and position occurred with equal probability. Each stimulus was presented for a duration of 500 milliseconds (ms), with an inter-stimulus interval of 2500ms.

Secondary task procedure. Relevant instructions, 30 practice trials, and stimuli for the *n*-back tasks were administered via E-Prime 2.08 (Schneider, Eschman, & Zuccolotto, 2002). During the task, 140 trials were administered with matches in 30% of trials. The *n*-back task varied along two dimensions, task domain (verbal or visuospatial) and task load (0-back or 2-back), resulting in four *n*-back task conditions: (1) verbal, 0-

back, (2) verbal, 2-back, (3) visuospatial, 0-back, (4) visuospatial, 2-back. A schematic of the task is presented in Figure 2.1. For the verbal task conditions, participants were required to monitor for matches across letter name, regardless of letter case. For the visuospatial task conditions, participants were required to monitor for matches across stimuli position on the screen. In order to minimize opportunities for verbal recoding in the visuospatial conditions, the term "letter" was not used in the instructions, and all training examples used a red square to demonstrate spatial location. The corresponding load manipulations for the verbal and visuospatial *n*-back tasks included 0-back and 2-back tasks. In the 0-back conditions, participants were instructed to press the space bar if the current stimulus matched the target stimulus presented at the beginning the 7-minute exposure block. In 2-back conditions, participants were instructed to press the space bar if a stimulus on each trial matched the stimulus occurring two trials previously.

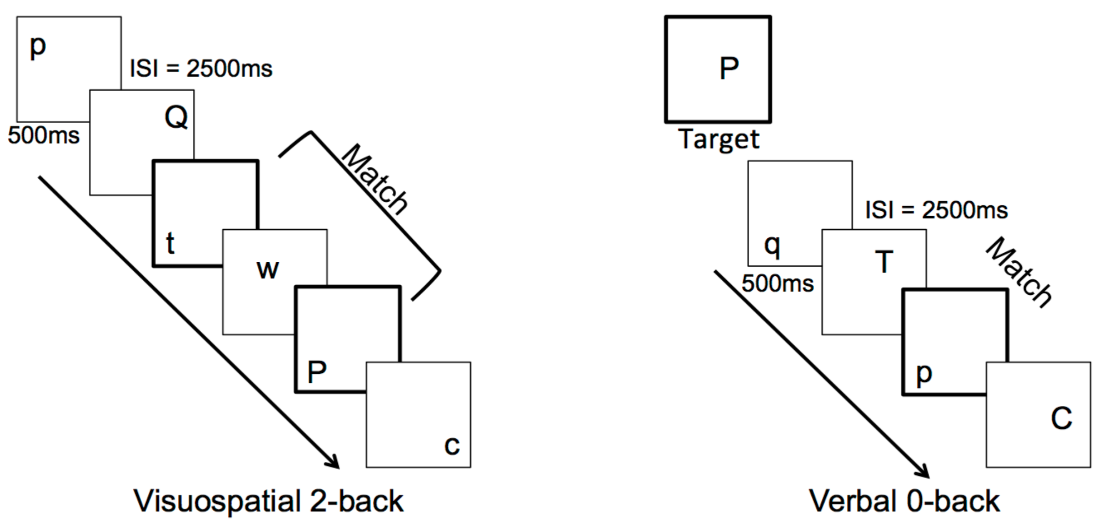


Figure 2.1 Schematic diagram of n-back task

2.2.2.3 Test phase

Test phase stimuli. Six non-word foils were constructed from the same 12 CV syllables as the artificial language: *pubati, tapudi, dupitu, tipabu, bidata, batipi*. Foil words were created with the constraint that within word transitional probabilities would be zero based on the participants' previous artificial language exposure. Syllables were drawn from the same recording inventory as the artificial language stimuli, with appropriate co-articulation contexts. Participants were tested on non-word foils rather than part-word foils, which consist of two syllables from a trained "word" from the artificial language, plus an incorrect syllable. Discrimination accuracy is generally higher for non-words than part-words (Saffran, Aslin, & Newport, 1996). As the aim of the present study was not to discover if statistical learning of language was present generally, but rather, whether it differed across experimental groups, the more sensitive measure was used.

Test phase procedure. Following the listening phase, participants completed a two-alternative forced-choice (2AFC) task delivered by E-Prime 2.08 (Schneider et al., 2002). For each test item, participants heard two tri-syllabic strings separated by 500ms of silence. One of these strings was a trained word from the artificial language, and the other a foil word. Subjects were instructed to indicate which word "sounds more like something you heard in the language", and to select "A" or "L" on the keyboard to indicate the first or second stimulus, respectively. The instructions stayed on the screen for the duration of the test phase. Each trained word was paired exhaustively with foil word, comprising 36 total test pairs, and pairs were presented in a fixed random order.

2.3 Results

2.3.1 Preliminary Analyses: n-back task performance

First, a preliminary analysis was conducted to ensure that the secondary n-back tasks imposed differing cognitive loads based on the recalling items 2 vs. 0 back, but not on task domain. Performance on the *n*-back was measured using a *d'* score, which accounts for both correct hits and false alarms (*see* Macmillan & Creelman, 1991). Mean *d'* scores for all *n*-back task groups are presented in Table 2.1. *N*-back performance data was missing from 4 participants in the verbal 0-back condition due to a computer error. Scores across secondary tasks were compared using a 2 (*domain*: verbal, visuospatial) by 2 (*load*: 0-back, 2-back) analysis of variance (ANOVA). There was a significant main effect of task load, F (1, 98) = 65.84, $p < .001$, β = 1.00, such that those in a 2-back condition had significantly lower *d*' scores on the *n*-back task than those in a 0-back condition. There was no significant main effect of task domain, F (1, 98) = 0.00, $p = .994$, $\eta_p^2 = .418$ and no significant interaction between task load and domain, F (1, 94) = 0.36, $p = .549$, $\eta_p^2 = .004$. These results confirmed that the 2-back conditions were more difficult than the 0-back conditions, given the lower *d'* scores across the 2-back conditions, and that performance on the *n*-back task did not differ as a function of task domain. It is important to note that even though the same stimuli were used across all task conditions, participants were sensitive to their target manipulation regarding task domain. This interpretation applies specifically to those in 0-back conditions, who were more accurate on the *n*-back task overall. That is, participants in the visuospatial conditions responded accurately to stimulus position, and participants in the verbal conditions responded accurately to letter name. This indicates that the participants were engaged in the domain-relevant task, which was crucial for the interference manipulation.

Table 2.1 Average d' scores on secondary n-back tasks across conditions.

	Domain	
Load	Verbal	Visuospatial
0-back	5.14 (2.66)	4.90 (1.55)
2-back	1.64 (2.11)	1.89 (1.48)

Note. Standard deviation in parentheses

2.3.2 Two-alternative forced-choice (2AFC) test

Table 2.2 shows the mean scores on the 2AFC task for all conditions. Performance on the 2AFC was compared across groups to determine whether engagement in the secondary tasks differentially impacted scores on the 2AFC test following language exposure. Because of the inclusion of the control (no-domain/no-load) condition in the analyses, the design was not fully factorial. Thus, a one-way analysis of variance (ANOVA) directly comparing across all task conditions was used. A significant effect of task condition on 2AFC scores was found, F (4, 126) = 3.791, p = .006, β = 0.881. Post-hoc comparisons using a Bonferroni correction revealed that those engaged either in the visuospatial 2-back, t (47) = 2.884, p = .005, d = 0.809, and verbal 2-back, t (49) = 3.195, p = .002, d = 0.886, had significantly lower scores than controls on the 2AFC measure. No other comparisons were significant.

Table 2.2 Average test scores for the 2-alternative forced choice task on word identification across secondary task conditions.

Secondary Task	*n*	*M*	*SD*
Control (no task)	26	22.27	3.67
Verbal 0-back	26	20.81	3.50
Verbal 2-back	26	18.81*^	4.13
Visuospatial 0-back	23	21.26	3.53
Visuospatial 2-back	25	19.12*^	4.10

Note. Scores are out of 36. * = group scores that were not statistically different from chance; ^ = group scores significantly lower than controls.

Next, single-sample t-tests were conducted to assess whether scores on the 2AFC task were significantly above chance across all conditions. Chance was defined as 18 items correct on the 2AFC, out of a possible total of 36; that is, 50% correct. Not all groups performed significantly better than chance: Scores for those in the visuospatial 2-back, $t\ (24) = 1.37$, $p = .092$, $d = 0.273$, and verbal 2-back tasks, $t\ (25) = 1.00$, $p = .164$, $d = 0.196$, did not differ from chance on the 2AFC measure, whereas the control task, $t\ (25) = 5.93$, $p < .001$, $d = 1.163$, visuospatial 0-back, $t\ (22) = 4.43$, $p < .001$, $d = 0.924$, and verbal 0-back, $t\ (25) = 4.09$, $p < .001$, $d = 0.803$, all performed above chance. A distribution of individual scores by secondary task group is presented in Figure 2.2.

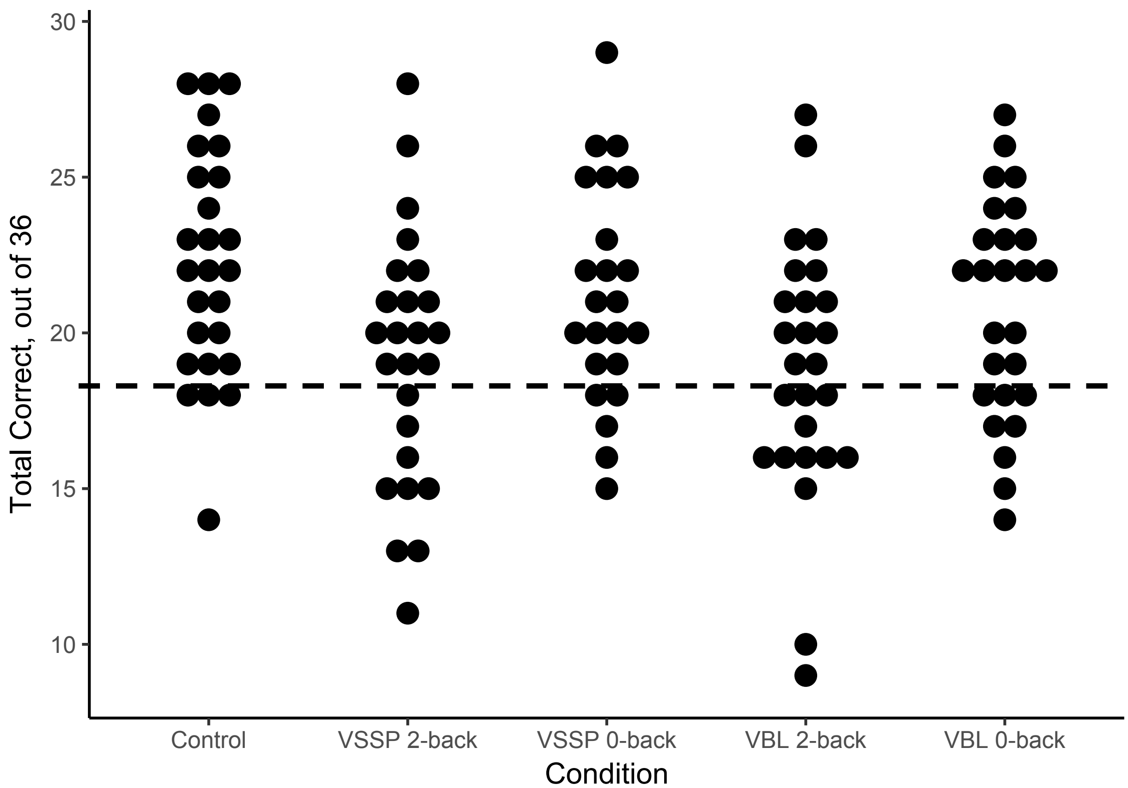


Figure 2.2 Scores on the two-alternative forced-choice task assessing word identification across secondary task groups. Horizontal dashed line represents chance responding (18 out of 36 correct responses).

Note. VSSP = visuospatial; VBL = verbal.

Taken together, these results clearly show that those completing a secondary 2-back task, regardless of task domain, had lower word identification scores compared to those completing no secondary task. These results are corroborated with the analyses comparing group-level performance to chance, wherein those completing a secondary verbal or visuospatial 2-back task did not score above chance on the 2AFC measure. On the other hand, those completing a verbal or visuospatial 0-back task identified words at an above-chance level following language exposure, and performed equivalently with those completing no secondary task.

2.4 Discussion

Studies examining interference effects on statistical learning have demonstrated both domain-general and domain-specific interference effects. However, because there is a wide diversity in the interference tasks used in these studies, and because of this, it difficult to meaningfully compare the interference effects. Because of this, it has been difficult to assess the whether there are true domain-general or domain-specific effects, and thereby, clearly elucidate the underlying cognitive mechanisms that may support statistical language learning. In the present study, the interference effects between domain-specific (verbal) and domain-general (visuospatial) tasks on auditory statistical word segmentation were compared. Participants were exposed to a structured, unsegmented artificial language while concurrently engaged in one of four secondary *n*-back tasks: The *n*-back task was either verbal (same-domain) or visuospatial (cross-domain), and involved either a 0-back or 2-back manipulation for either task domain. Following language exposure, scores on a word identification task were compared across secondary task groups relative to a control (no secondary task) condition. It was found that those completing a secondary 2-back task, regardless of task domain, had lower word identification scores than controls. In fact, those completing either the verbal or visuospatial 2-back task did not score above chance on the word identification task, suggesting that they did not successfully segment the words from the artificial language. All other groups performed above chance.

Overall, this pattern of results demonstrates an interference effect on statistical language learning from a secondary task spanning both verbal and non-verbal domains, and is consistent with theories suggesting that statistical language learning is supported by domain-general cognitive resources. One theory relevant to these findings is the framework for statistical learning proposed by Frost and colleagues (2015), which suggests that domain-general neurobiological mechanisms and computational principles operate in different modalities, and learning is constrained by the modality of the incoming stimulus. Thus, partially shared domain-general neural networks are responsible for the computation of to-be-learned stimuli. These on-line computations of transitional probabilities may have been disrupted by the cross-domain 2-back tasks. Given this, it may be possible to infer that domain-general processing resources, which support shared domain-general computational principles, are necessary to support statistical language learning and indeed contribute to statistical language learning. It is important to note that the secondary same-domain verbal task (letter monitoring) was not of the same modality as the auditory statistical learning task. Frost et al.'s (2015) model, however, allows for shared processing of cross-modality audio-visual verbal stimuli. The results are in line with shared domain-specific but not modality-specific processing of audio and visual verbal stimuli, perhaps within a shared phonological resource within the language domain. In future research, it may be necessary to consider modality-specific, domain-specific, and domain-general constraints in order to understand the encoding and extracting of statistical regularities, at least with regards to phonological information.

The effect of domain-general interference on statistical language learning observed here may be due to the engagement of partially shared resources modulating the encoding of both the auditory language sequence and the *n*-back task stimuli. Uncovering the putative nature of these shared cognitive resources has significant implications for theories of statistical language learning, and statistical learning more generally. One likely explanation is that statistical learning involves memory-related processes including encoding, storing, and accessing newly learned information at retrieval (Erikson & Thiessen, 2015; Thiessen, Kronstein, & Hufnagle, 2013). Memory-based perspectives on statistical learning would postulate that disruptions to statistical learning are due to memory interference or decay (Thiessen, 2017). Interference in the task would have been caused by the 2-back manipulations of the *n*-back task. The 2-back task involves constant updating and monitoring of incoming stimuli, resulting in the active engagement of participants' working memory. Working memory is a component of general memory processes (D'Esposito, 2007) and, importantly, is a capacity limited system (e.g., Baddeley & Hitch, 1974). Additionally, working memory has shown to be involved in word learning, including in the acquisition of new words (Baddeley, Gathercole, & Papagno, 1988; Gathercole, 2006; Gathercole & Baddeley, 1989). It follows, then, that if the capacity limit of working memory was perhaps met or exceeded by engagement in the 2-back tasks, the necessary additional resources required for segmenting the language stream would not have been available.

The suggestion that statistical word segmentation is disrupted by concurrent working memory engagement is analogous with findings from Palmer and Mattys (2016), who found statistical learning interference effects under verbal and visuospatial 2-back concurrent task conditions. While this collection of results provides evidence of interference in statistical learning caused by concurrent working memory engagement, the nature of this interference needs to be elucidated. One possibility is that due to the concurrent working memory task, listeners are unable to hold the complete phonological form of a newly segmented word in auditory working memory, resulting in mis-segmentations during language learning. One of the key stipulations in PARSER (Perruchet & Vinter, 1998), a memory-based account of statistical learning, is that syllable sequences are refreshed or reactivated upon repeated occurrence. Refreshing the newly segmented syllable sequence leads to integration across repeated presentations and, subsequently, the retention of the candidate word within memory. If the auditory stream is mis-segmented, an accurate match of the extracted candidate word may not be re-encountered with the precision necessary to integrate across repeated presentations, nor to store the word within memory. Further exploration of the nature of these mis-segmentations may be necessary.

There are two possible findings that could be predicted from a memory-based model of statistical learning. First, it is possible that participants whose statistical learning was disrupted due to memory limitations may be familiar with part-words, that is, two adjacent syllables from one of the tri-syllabic trained words. This familiarity would be due to mis-segmentations of the artificial language stream due to limitations within the phonological store. Second, participants' familiarity with the words may be graded as a function of word frequency. Words that were encountered more frequently may have been sufficiently refreshed within memory, and thereby stored as candidate words, despite memory limitations. Future research would be needed to investigate these possibilities.

It is interesting to speculate as to why the domain-general interference effects in the present study, with 7 minutes of language exposure, are inconsistent with the domain-specific interference effect found in a previous study (Noonan, 2014), which featured 28 minutes of language exposure. When considering these effects, is important to note that the analysis of performance on the *n*-back task showed equivalent performance across task domains (verbal, visuospatial), indicating that they were of equal difficulty and, presumably, imposed similar levels of interference. Toro and colleagues (2005) demonstrated that word segmentation performance dropped when attention was diverted to a different stream within the same sensory modality (auditory), or to a different modality (vision). Costs to statistical learning when the concurrent task was of a different modality only occurred when the distractor task was more difficult, whereas costs to statistical learning for the auditory task were consistent across task difficulties. It is possible to interpret the results from the present study in light of these findings. First, the present findings of interference effects for the visuospatial 2-back condition, but not the 0-back condition is consistent with the conclusion that it is more difficult to find attention costs across modalities than within a modality (Duncan, Martens, & Ward, 1997; Soto-Faraco & Spence, 2002; Triesman & Davies, 1973; Toro et al., 2005). However, this conclusion is weakened by the fact that there were no interference effects for the verbal 0-back condition, wherein domain-specific attention costs across task difficulty levels could have been expected.

Another possibility is that there are variations in individual differences in statistical learning across the samples used in the present study and in Noonan (2014), which the present analyses were unable to capture. Prior work has suggested that there is wide and measurable individual variation in statistical learning (Arciuli & Simpson, 2012; Bogaerts, Siegelman, & Frost, 2016; Conway et al., 2010; Evans et al., 2009; Karuza et al., 2013; Misyak & Christiansen, 2012; Siegelman & Frost, 2015). However, the statistical learning assessment used in the present study, the two-alternative forced-choice task (2AFC), is not suitable to analyze individual performance data, and may not be a reliable or valid measure of statistical learning (Siegelman, Bogaerts & Frost, 2016). Given these shortcomings, the increased use of implicit measures of statistical language learning, such as the measurement of neural responses (Alba, Katahira, & Okanoya, 2006; Batternik, Reber, Neville, & Paller, 2015; Cunillera, Toro, Sebastián-Gallés, & Rodríguez-Fornells, 2006; Sanders, Newport, Neville, 2002), may be informative in understanding the individual variations in sensitivity to statistical structures.

2.4.1 Conclusions

It has been proposed that first language acquisition occurs, in part, via an implicit statistical learning mechanism, and this mechanism may be specifically useful for helping language learners discover the boundaries between words in fluent speech. What has been poorly understood, however, are the cognitive processes that support implicit statistical language learning. To address this, this study sought examined how engaging in explicit same-domain (verbal) or cross-domain (visuospatial) cognitive tasks influenced implicit statistical learning of word boundaries in an artificial language. It was found that engagement in a secondary 2-back task, regardless of task domain, resulted in interference in statistical language learning, while engagement in a 0-back verbal or visual task did not. This cross-domain interference effect supports the model that statistical learning is supported by domain-general cognitive resources rather than representing a strictly modality-specific learning mechanism.

2.5 References

Abla, D., Katahira, K., & Okanoya, K. (2008). On-line assessment of statistical learning by event-related potentials. *Journal of Cognitive Neuroscience*, *20*(6), 952-964.

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

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9,* 321-324.

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology, 63*, 1-29.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of Learning and Motivation*, *8*, (47-89).

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language*, *83*, 62-78.

Bogaerts, L., Siegelman, N., & Frost, R. (2016). Splitting the variance of statistical learning performance: A parametric investigation of exposure duration and transitional probabilities. *Psychonomic Bulletin & Review,* 1-7.

Chen, L., & Vroomen, J. (2013). Intersensory binding across space and time: A tutorial review. *Attention, Perception, & Psychophysics, 75,* 790-811.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT press.

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

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

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.

Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition*, *114*(3), 356-371.

Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant melodies: statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, memory, and cognition*, *30*(5), 1119.

Cunillera, T., Toro, J. M., Sebastián-Gallés, N., & Rodríguez-Fornells, A. (2006). The effects of stress and statistical cues on continuous speech segmentation: an event-related brain potential study. *Brain Research*, *1123*(1), 168-178.

D'Esposito, M. (2007). From cognitive to neural models of working memory. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *362*(1481), 761-772.

Dienes, Z., Altmann, G., Kwan, L., & Goode, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(5), 1322.

Duncan, J., Martens, S., & Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature*, *387*(6635), 808.

Erickson, L. C., & Thiessen, E. D. (2015). Statistical learning of language: theory, validity, and predictions of a statistical learning account of language acquisition. *Developmental Review, 37,* 66-108.
Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 52,* 321-335.
Fernandes, T., Kolinsky, R., & Ventura, P. (2010). The impact of attention load on the use of statistical information and co-articulation as speech segmentation cues. *Attention, Perception, & Psychophysics, 72*(6), 1522-1532.
Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science, 12*, 499-504.
French, R. M., Addyman, C., & Mareschal, D. (2011). TRACX: A recognition-based connectionist framework for sequence segmentation and chunk extraction. *Psychological Review*, *118*(4), 614.
Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: the paradox of statistical learning. *Trends in Cognitive Sciences, 19,* 117-125.
Gathercole, S. E. (2006). Non-word repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*(4), 513-543.
Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, *28*(2), 200-213.
Gebhart, A. L., Newport, E. L., & Aslin, R. N. (2009). Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. *Psychonomic Bulletin & Review*, *16*(3), 486-490.
Gleitman, L. R., & Wanner, E. (1982). *Language Acquisition: The State of the Art*. CUP Archive.
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., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, *4*(5), 178-186.
Graf-Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science, 18,* 254-260.
Johansson, T. (2009). Strengthening the case for stimulus-specificity in artificial grammar learning: no evidence for abstract representations with extended exposure. *Experimental Psychology, 56*, 188-197.
Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., & Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain and Language*, *127*(1), 46-54.
Keele, S. W., Ivry, R., Mayr, U., Hazeltine, E., & Heuer, H. (2003). The cognitive and neural architecture of sequence representation. *Psychological Review*, *110*(2), 316.
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.
Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews: Neuroscience, 5,* 831-843.

Lew-Williams, C., Pelucchi, B., & Saffran, J. R. (2011). Isolated words enhance statistical language learning in infancy. *Developmental Science, 14,* 1323-1329.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection Theory: A User's Guide*. New York, NY: Cambridge University Press.

Mareschal, D., & French, R. M. (2017). TRACX2: a connectionist autoencoder using graded chunks to model infant visual statistical learning. *Philosophical Transactions of the Royal Society of London B*, *372*(1711), 20160057.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82,* B101-B111.

Mayr, U. (1996). Spatial attention and implicit sequence learning: Evidence for independent learning of spatial and nonspatial sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(2), 350.

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: An individual differences study. *Language Learning*, *62*(1), 302-331.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, *14*(1), 11-28.

Noonan, N. B. (2014). The relationship between implicit and explicit processing in statistical language learning. (Unpublished master's thesis). The University of Western Ontario, London, Ontario.

Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language, 55*(4), 601-626.

Palmer, S. D., & Mattys, S. L. (2016). Speech segmentation by statistical learning is supported by domain-general processes within working memory. *The Quarterly Journal of Experimental Psychology, 69*(12), 2390-2401.

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Learning in reverse: Eight-month-old infants track backward transitional probabilities. *Cognition, 113*, 244-247.

Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science, 298*(5593), 604-607.

Perruchet, P., & Vinter, A. (1998). PARSER: A model for word segmentation. *Journal of Memory and Language*, *39*(2), 246-263.

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

Pioneer Hill Software. (2008). *Spectraplus* (Version 5.0.26.0). Poulsbo, Washington: Pioneer Hill.

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

Recanzone, G. H. (2009). Interactions of auditory and visual stimuli in space and time. *Hearing Research, 258*, 89-99.

Romberg, A. R., & Saffran, J. R. (2013). All together now: Concurrent learning of multiple structures in an artificial language. *Cognitive Science*, *37*(7), 1290-1320.

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

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.

Salminen, N. H., Aho, J., & Sams, M. (2013). Visual task enhances spatial selectivity in the human auditory cortex. *Frontiers in Neuroscience, 7*, 44.

Sanders, L. D., Newport, E. L., & Neville, H. J. (2002). Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. *Nature Neuroscience*, *5*(7), 700.

Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M., & Turk-Browne, N. B. (2014). The necessity of the medial temporal lobe for statistical learning. *Journal of Cognitive Neuroscience, 26*, 1736-1747.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime 1.0. Pittsburgh, PA: Psychological Software Tools.

Seitz, A. R., Kim, R., van Wassenhove, V., & Shams, L. (2007). Simultaneous and independent acquisition of multisensory and unisensory associations. *Perception, 36,* 1445-1453.

Shanks, D. R., Johnstone, T., & Le Staggs, A. (1997). Abstraction processes in artificial grammar learning. *The Quarterly Journal of Experimental Psychology: Section A, 50*, 216-252.

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language, 81*, 105-120.

Siegelman, N., Boegarts, L., & Frost R. (2016). Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behavioural Research Methods*, 1-15.

Soto-Faraco, S., & Spence, C. (2002). Modality-specific auditory and visual temporal processing deficits. *The Quarterly Journal of Experimental Psychology Section A*, *55*(1), 23-40.

Sound Forge Audio Studio (2013). *Sony Creative Software Inc*.

SpectraPlus software (2008). *Pioneer Hill Software.*

Teinonen, T., Fellman, V., Näätänen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neuroscience, 10*, 21.

Thiessen, E. D. (2017). What's statistical about learning? Insights from modelling statistical learning as a set of memory processes. *Philosophical Transactions of the Royal Society of London B*, *372*(1711), 20160056.

Thiessen, E. D., & Pavlik Jr, P. I. (2013). iMinerva: a mathematical model of distributional statistical learning. *Cognitive Science*, *37*(2), 310-343.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, *7*(1), 53-71.

Thiessen, E. D., Kronstein, A. T., & Hufnagle, D. G. (2013). The extraction and integration framework: A two-process account of statistical learning. *Psychological Bulletin, 139*, 792-814.

Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition, 97*(2), B25-B34.

Treisman, A. M., & Davies, A. (1973). Divided attention to ear and eye. *Attention and performance IV*, 101-117.

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

USBPre 2 pre-amplifier. *Sound Devices LLC.*

Chapter 3: Examining the relationship between statistical learning and cognitive abilities in school-aged children

3.1 Introduction

Statistical learning is generally viewed as a domain-general mechanism that uncovers the distributional regularities across stimulus types (Reber, 1967; Saffran & Thiessen, 2007). But, there is a lack of clarity as to what is meant by "domain-generality". Some accounts ascribe domain generality in statistical learning to a unitary, general purpose learning mechanism (Bluf, Johnson, & Valenza, 2011), that can operate uniformly across domains (Saffran & Thiessen, 2007), and in non-human species (Hauser, Newport, & Aslin; 2001; Milne, Wilson, & Christiansen, 2018). Despite this, similar findings of statistical learning across domains does not necessarily implicate the same underlying cognitive mechanism. Indeed, there is evidence showing that statistical learning is modality- or stimulus-specific (Altmann, Dienes & Goode, 1995; Conway & Christiansen, 2005; Redington & Chater, 1996, *see* Frost, Armstrong, Siegelman, & Christiansen, 2015). Of potential interest to the question of domain-general vs. domain-specific effects in statistical learning would be an examination of performance of a group with a disproportionate language impairment such as children with a developmental language disorder (DLD). Previous work has revealed marked DLD deficits in statistical language learning tasks (SLL) tasks (e.g., Evans, Saffran, & Robe-Torres, 2009; Mainela-Arnold & Evans, 2014) as well as non-linguistic sequential learning tasks (e.g., Lum, Conti-Ramsden, Morgan, & Ullman, 2014; Obeid, Brooks, Powers, Gillespie-Lynch & Lum, 2016), however no previous studies have directly compared statistical learning across verbal and visual domains in children with DLD. The purpose of the present study was to investigate whether statistical learning is subject to domain-specific processing constraints, or is perhaps subject to more domain-general limitations by comparing statistical learning across verbal and visual domains in children with either typical development or DLD, and to examine the relationship between statistical learning and other measures of language and cognition.

The nature of domain-generality in statistical learning has been linked to both common computational processes involved in statistical learning across domains (Thiessen & Saffran, 2007), as well as domain-general cognitive processes that may support statistical learning (Frost et al., 2015). In describing the domain-general computational processes involved in statistical learning, Frost et al. (2015) proposed a theoretical model that construed statistical learning as a set of domain-general neurobiological mechanisms for learning, detecting, and processing the distributional regularities across stimulus types. This learning system is not a unitary system, as processing within a domain is subject to the specific perceptual constraints of that domain. However, there are common domain-general computational principles that operate across domains, or partially shared neural networks that modulate the encoding of to-be-learned material (e.g., Fedorenko & Thomspon-Schill, 2014). Although not acknowledged in Frost et al.'s (2015) account, domain-general cognitive processes may also play a role in mediating statistical learning. Some work has demonstrated that statistical learning may depend on attention (Toro, Sinnett, & Soto-Faraco, 2005, *c.f.* Saffran, Newport, Aslin, Tunick, & Barrueco, 1997).

Others have shown that working memory may also be involved in SLL in a domain-general way (Palmer & Mattys, 2016; *see* Chapter 2 *of this thesis*). In terms of SLL, working memory may play a role as an active maintenance mechanism for newly segmented words via the phonological loop component of working memory (*see* Lopez-

Barroso, de Diego-Balaguer, Cunillera, Camara, Münte, & Rodríguez-Fornells, 2011). Studies showing working memory interference in statistical language tasks (e.g., Palmer & Mattys, 2016; Lopez-Barroso et al., 2011; *see* Chapter 2 *of this thesis*), have raised the possibility that limitations in working memory lead to impairments in statistical learning. Working memory likely mediates statistical learning via short term or temporary storage of segmented units: In order to detect the distributional regularities within linguistic input to segment words from fluent speech, one must be able to retain the phonological sequence within memory. If the memory span is not sufficient to hold the sequence in mind, the regularities between adjacent elements cannot be detected. Initial evidence for this comes from studies showing impairments in statistical learning due to working memory interference (Palmer & Mattys, 2016; *see* Chapter 2 *of this thesis*). Consistent with this notion, a recent account of statistical learning suggests that statistical learning arises from a set of memory processes involving the extraction of an element of the input into a memory trace, and the subsequent integration across these stored memory traces to detect the distributional regularities (Thiessen, 2017). This memory-based account of statistical learning is linked theoretically to working memory accounts of vocabulary acquisition where word learning is mediated by temporary phonological storage in verbal short-term memory (Baddeley, Gathercole, & Papagno, 1998; Gathercole, 2006). These memory-based accounts of word learning leaves open the possibility that a verbal working memory limitation may lead to an impairment in SLL. The role of working memory in non-linguistic tasks, however, has not been previously examined. Findings of a deficit in statistical learning across different domains in children known to have a working memory impairment would provide additional evidence regarding the role of working memory in statistical learning. In order to examine this possibility, the present study investigated the relationships between verbal and visual statistical learning, and between these tasks and measures of verbal and visuospatial working memory.

Although domain-general processes are likely essential to successful statistical learning, there are nevertheless domain-specific constraints on statistical learning. Such constraints are explicitly acknowledged within the theoretical model proposed by Frost et al. (2015) due to perceptual processing differences across domains, and are supported by empirical findings. For example, Conway and Christiansen (2005) demonstrated both quantitative and qualitative differences in statistical learning of structured material across verbal, visual, and tactile domains. Domain-specificity in statistical learning is also supported by neuroimaging research. Recent fMRI studies have shown that VSL results in activation in visual regions, including lateral occipital cortex and inferior frontal gyrus (Turk-Browne, Scholl, Chun, & Johnson, 2009), while auditory statistical learning invokes activation in auditory areas, including left middle and frontal gyri (McNealy, Mazziotta, & Dapretto, 2006), as well as the posterior superior temporal gyrus (Cunillera et al., 2009). In examining the domain-specificity of statistical learning, there are inherent challenges in designing similar tasks that differ only in the domain of the material. Notably, this was addressed by Conway and Christiansen (2005), who sought to use closely comparable cross-domain tasks. Similarly, in the present study, both the SLL and VSL tasks involved the computation of transitional probabilities in order to segment either tri-syllabic or three-shape sequences. Findings that children with a disproportionate impairment in one domain, such as language, are impaired on an SLL task but not a VSL task would provide additional evidence in support of domain-specific constraints on statistical learning, as well as findings of correlations between SLL and language measures, but not domain-general cognitive measures.

Before proceeding, it is important to acknowledge the relevance of procedural learning, and how it can inform hypothesis related to statistical learning, as much of the work examining implicit learning in language-impaired groups has used procedural learning tasks (*see* Lum et al., 2014). There has been some discussion in the field of implicit learning on the distinction between procedural and statistical learning, as these terms are sometimes used interchangeable (Perruchet & Pacton, 2006). The term "procedural learning" is most consistently applied to implicit learning tasks, which focus heavily on the unconscious nature of learning in these paradigms (Reber, 1967). Common examples of such tasks include artificial or finite-state grammar tasks and serial reaction time (SRT) tasks, both of which involve the gradual learning of structured material. "Statistical learning", when it first originated, was used to describe the ability of infants to discover words in a continuous speech stream (Saffran, Aslin, & Newport, 1996), and has since been applied to learning or segmenting stimuli across a variety of stimulus types (e.g., Conway & Christiansen, 2005). With regards to language learning, these terms were contrasted in that implicit or procedural learning pertained to syntactic acquisition, or the learning of rule-like structures (e.g., Marcus, Vijayan, Rao, & Vishton, 1999), while statistical learning dealt with lexicon formation (e.g., Saffran, 2003). However, statistical learning is no longer simply applied to learning word-like units, but has been shown to operate in learning complex structures such as syntax (Gómez, 2002, Saffran & Wilson, 2003). One theoretical distinction that remains is that implicit learning tends to favour the formation of chunks (e.g.: Servan-Schreiber & Anderson, 1990), while statistical learning tends to favour statistical computations (Aslin, Saffran, & Newport, 1998; Saffran, 2001). Both of these operations have been argued to be independent processes (Meulmans & van der Linden, 2003; *see* Perruchet & Pacton, 2006 *for a discussion*). However, others have argued that chunk formation is the default process, with sensitivity to statistics being a by-product of this process (Perruchet & Vinter, 1998). Further still, a recent discussion (Christiansen, 2018) has sought to align implicit and statistical learning under the umbrella term *implicit statistical learning*, which reconciles much of the related findings between these two bodies of literature. For the purposes of the present study, findings from procedural and statistical learning paradigms will be discussed. This collection of findings can, at minimum, be closely aligned, with findings from procedural learning paradigms being useful when forming hypotheses related to statistical learning.

Given the growing research interest in exploring the domain general and domain specific constraints on statistical language learning, it is of interest to examine these constraints in a clinical population with a disproportionate deficit in one domain. To do this, the present study explored statistical learning across verbal and visual domains via SLL and VSL tasks in children with a persistent language disorder not associated with a known biological etiology known as developmental language disorder (DLD; Bishop, Snowling, Thompson, Greenhalgh, & CATALISE-2, 2017). DLD is a relatively common developmental disorder, affecting roughly 7% of kindergarten children (Norbury et al., 2016; Tomblin, Records, Buckwalter, Zhang, Smith, & O'Brien, 1997), and is more prevalent in male than female children (e.g., Choudhury & Benasich, 2003; Flax, Raelpe-Bonilla, Hirsch, Brzustowicz, Bartlett, & Tallal, 2003). These children show difficulties in learning the grammatical structures of language, with the acquisition of pragmatics and semantics relatively intact (Leonard, 1998). Language abilities are markedly impaired in DLD, as evidenced by poor performance relative to typically developing peers on a variety of linguistic measures, including tasks of grammaticality judgment (Miller, Leonard, & Finneran, 2008, Rice, Wexler, & Redmond, 1999), sentence comprehension

(Bishop, 1979, 1982, 1997), sentence formulation (Leonard, Miller, Grela, Holland, Gerber, & Petucci, 2000), and object naming (McGregor, Newman, Reilly & Capone, 2000). Further, impairments in verbal working memory are common for those with DLD, with visuospatial working memory relatively spared (*see* Archibald & Gathercole, 2006a). These clear linguistic deficits make DLD a useful and informative clinical case to examine the relationship between language and statistical learning across domains.

In one of the first investigations of statistical learning in a DLD sample, Evans, Saffran, and Robe-Torres (2009) compared performance across both a statistical word segmentation task and a tone-sequence segmentation task. The authors speculated that children with DLD would show a statistical learning impairment, and that this impairment would be related to their language difficulties. The statistical word segmentation task was similar to the one used by Saffran, Newport, and Aslin (1996), wherein tri-syllabic words could only be segmented from a speech stream based on the transitional probabilities between adjacent syllables. The tone-sequence stimuli were structured similarly to the word segmentation task (*see* Saffran, Johnson, Aslin, & Newport, 1999). In this task, pure tones were played one at a time in a structured, unsegmented stream, and three-tone sequences could only be segmented based on the transitional probabilities between adjacent tones. Participants were exposed to the language stream for 21 minutes (Experiment 1) and 42 minutes (Experiment 2a), and the tone stream for 42 minutes (Experiment 2b). The DLD group did not perform above chance on the word segmentation task with 21 minutes of exposure, however, their performance improved to above-chance levels when the exposure duration was increased to 42 minutes. Typically developing (TD) children, on the other hand, consistently performed above chance across both exposure durations. In investigating the relationship between SLL and language abilities, performance on the word segmentation task at 21 minutes of exposure was found to be related to receptive and expressive vocabulary scores for the DLD group. For the tone-sequence learning task, the TD group performed above chance while the DLD group did not, despite the lengthy exposure period. Given this diverging performance for the DLD group across the two statistical learning tasks, Evans et al. (2009) speculated that statistical learning is not a domain-general mechanism in this population. However, without assessing statistical learning in a non-auditory task, it is difficult to determine the domain-specificity of the statistical learning impairment in those with DLD.

In a subsequent DLD study, Mainela-Arnold and Evans (2014) examined the relationship between SLL and both a lexical-phonological and a lexical-semantic task. Similar to the Evans et al. (2009) study, children with DLD completed a statistical word segmentation task (e.g., Saffran, Newport & Aslin, 1996). Mainela-Arnold and Evans (2014) found a DLD deficit on the statistical word segmentation task, and that this deficit was related to poor performance on the lexical-phonological task, but not the lexical-semantic task. Taken together with the association between statistical word segmentation and vocabulary found in Evans et al. (2009), Mainela-Arnold and Evans (2014) suggested that a procedural learning deficit, as captured by poor performance on the statistical word segmentation task, was related to the poor lexical-phonological abilities in DLD. The poor lexical-phonological abilities, then, resulted in the poor vocabulary outcomes for the DLD group. However, as with the Evans et al. (2009) study, it is difficult to determine if the finding of impaired statistical learning in DLD from Mainela-Arnold and Evans (2014) is due to a domain-specific or domain-general statistical learning impairment in this population. Or, whether statistical learning is related to other cognitive abilities not measured in the above studies. In order to expand this research area and determine whether statistical learning is impaired generally in this population, a more comprehensive approach is needed.

Departing from statistical word segmentation tasks, Mayor-Dubois, Zesiger, van der Linden, and Roulet-Perez (2014) examined procedural learning in children with DLD across a variety of tasks to further examine the domain-specificity of a statistical or procedural learning deficit in this group. Participants completed a phonotactic learning task, a motor sequence learning task, and a cognitive learning task. The phonotactic learning task resembled that of Majerus, van der Linden, Mulder, Meulemans, and Peters (2004), wherein participants heard a continuous sequence of 3000 consonant-vowel syllables. Artificial phonotactic rules determined the possible combinations between phonemes and syllables. Following exposure to the syllable sequence, participants completed a "lexical" decision task similar to the 2AFC used in statistical word segmentation studies (e.g., Saffran, Newport & Aslin, 1996). Consistent with the findings from Evans et al. (2009) and Mainela-Arnold and Evans (2014), Mayor-Dubois et al. (2014) found that children with DLD were unable to learn the phonotactic sequences providing further evidence in favour of a procedural learning impairment in DLD. Moreover, Mayor-Dubois et al. (2014) did not find an impairment for the DLD group on the motor sequence learning task nor the cognitive learning task, suggesting that the DLD deficit in statistical or procedural learning is specific to learning linguistic sequences.

A number of recent meta-analyses have provided additional evidence of a procedural or statistical learning deficit in DLD. What is significant from these analyses is that they provide evidence of procedural learning impairments for those with DLD on non-linguistic tasks, supporting the notion that this group is impaired on domain-general sequence learning tasks. In the first of such analyses, Lum and colleagues (2014) examined Serial Reaction Time (SRT) task performance in children with DLD across eight studies. The SRT task is a parameter commonly used to assess implicit sequence learning (Nissen & Bullemer, 1987). For this task, participants make a speeded response to a fixed set of stimuli. Unbeknownst to the participant, there are probabilities governing the transition between cues. Sequence learning on this task is defined as a decrease in reaction times (i.e., faster responding) as subjects learn the transitional probabilities between stimuli. Lum et al. (2014) found SRT impairments in DLD compared to typically developing peers, pointing to a procedural learning deficit in this population. In a subsequent study, Obeid and colleagues (2016) examined procedural and statistical learning across a variety of tasks in individuals with DLD. In their analysis, they included studies examining SRT, contextual cuing, artificial grammar learning, speech segmentation, observational learning, and probabilistic classification. They found that the DLD group was impaired across procedural and statistical learning tasks, and the effect size related to this deficit did not differ due to task modality. Specifically examining auditory statistical learning in DLD, Lammertink, Boersma, Wijnen, and Rispens (2017) found a robust DLD deficit for the detection of statistical regularities in auditory input for both verbal and non-verbal auditory stimuli. Taken together, there is mounting evidence for a procedural or statistical learning deficit in those with DLD, and that this deficit may be domain-general.

The domain-general deficits seen on statistical or procedural learning tasks in DLD are difficult to reconcile with the literature supporting the relative linguistic deficit in this group. In order to closely examine whether children with DLD are impaired on only a verbal statistical learning task, or whether they are also impaired on a non-verbal task, the present study also included measures of both SLL and VSL. Participants in the present study completed a statistical word segmentation task similar to the task in Saffran, Newport, Aslin, Tunick, and Barrueco (1997), which has been used in previous statistical learning studies with DLD groups (e.g., Evans et al., 2009; Mainela-Arnold & Evans, 2014). The VSL task was a close replication of the task developed by Siegelman, Bogaerts, and Frost (2016), adapted for use in a developmental sample. Similar to a word segmentation task, this VSL task involves the presentation of a structured, unsegmented stream wherein shapes are organized into triplets based on the transitional probabilities between adjacent shapes. As with the SLL task, within-triplet transitional probabilities were higher than between-triplet transitional probabilities (0.5-1.0 within-triplet *versus* > 0.3 between-triplet). The test for the VSL task incorporated both pattern recognition items and pattern completion items. The inclusion of these test items both maximize the number and diversity of questions asked, while circumventing some of the psychometric limitations of only using a 2AFC test (*see* Siegelman et al., 2016 *for a discussion*). A foundational study on VSL showed that adults can learn the relationships embedded in sequences of shapes after only 6 minutes of exposure (Fiser & Aslin, 2002, *also see* Turk-Browne et al., 2005), and similar tasks have demonstrated successful statistical learning of visual sequences in typically-developing school-aged samples (Arciuli & Simpson, 2011). Examining VSL in children with DLD provides a useful analogue to a statistical word segmentation task, and adds substantially to the empirical evidence on domain-general statistical learning in this population.

Although predicting an impairment on SLL for those with DLD seems to be an obvious extension of previous findings (e.g., Evans et al., 2009; Mainela-Arnold & Evans, 2014), predicting their performance on a VSL is somewhat less clear. One possible theory that would predict a domain-general statistical or procedural learning deficit in DLD, and the interpretation favoured by earlier research (e.g., Evans et al., 2009; Lum et al., 2014; Mainela-Arnold, 2014), is the Procedural Deficit Hypothesis. The Procedural Deficit Hypothesis states that a procedural memory impairment underlies the language difficulties seen in children with DLD, specifically those difficulties associated with grammatical processing. Originally proposed by Ullman and colleagues (Ullman, 2001; Ullman & Gopnik, 1999; Ullman & Pierpont, 2005), the Procedural Deficit Hypothesis suggests that there is a distinction between lexical and grammatical knowledge. The declarative system is responsible for the acquisition and representation of the mental lexicon, specifically the form-meaning associations related to word-specific knowledge. The procedural learning system, on the other hand, is thought to support the acquisition of the sequential representations that are characteristic of grammatical knowledge and other types of domain-general procedural learning. Thus, this theory would predict DLD impairments in procedural or statistical learning tasks, and that these difficulties should be related to the grammar difficulties seen in this population as grammar learning is mediated by the procedural learning system. Some evidence showing poor procedural learning of novel grammatical strings in adolescents with DLD supports this conclusion (Tomblin, Mainela-Arnold, & Zhang, 2007). Additionally, the Procedural Deficit Hypothesis can account for poor procedural or statistical learning in DLD groups in non-linguistic domains, as it is a domain-general hypothesis (Ullman & Pierpont, 2005). What is difficult to reconcile with the Procedural Deficit Hypothesis are the findings from Evans et al. (2009) and Mainela-Arnold and Evans (2014), where poor procedural learning was associated with poor vocabulary in the former study, and poor lexico-phonological learning in the latter study, as both of these are generally assumed to hinge on associative rather than procedural mechanisms. A recent meta-analysis

(Hamrick, Lum & Ullman, 2018) concluded that lexical abilities are exclusively mediated by the declarative system, while grammar was linked to both the declarative and procedural learning systems. This conclusion makes it difficult to discern how an association between statistical learning and vocabulary outcomes can be reconciled with the Procedural Deficit Hypothesis, which posits a strict distinction between lexical and grammatical learning.

Given the theoretical link between working memory and statistical learning, it is possible that if those with DLD have an insufficient working memory capacity to hold a newly-segmented word as a phonological trace, or to properly segment the speech stream, this would manifest in impairments on SLL outcome measures. There is a considerable body of evidence that children with DLD often present with comorbid working memory deficits, specifically in the verbal domain. Children with DLD tend to be impaired on tasks of verbal short-term and working memory (Archibald & Gathercole, 2006b; Marton, Eichorn, Campanelli, & Zakarias, 2016; Montgomery, 2003). The finding of a verbal working memory deficit in children with DLD is particularly pronounced on tasks involving the immediate recall of unfamiliar phonological forms, or non-words (Archibald & Joanisse, 2009; Bishop, Bishop, North, & Donlan 1996; Coady & Evans, 2008; Conti- Ramsden, 2003; Conti-Ramsden, Botting, & Faragher, 2001; Dollaghan & Campbell, 1998; Edwards & Lahey, 1998; Weismer, Tomblin, Zhang, Buckwalter, Chynoweth, & Jones, 2000; Montgomery, 1995). In fact, investigations of working memory impairments in those with DLD have identified subtypes of this disorder whose language difficulties may be characterized by dual deficits in language and working memory (Archibald & Joanisse, 2009; Noonan, Redmond, & Archibald 2014). Consistent with this evidence, Evans et al. (2009) did speculate that a working memory impairment could explain the poor statistical learning in their DLD group. In a comprehensive review, Hsu and Bishop (2011) hypothesized that the grammatical difficulties in those with DLD are a consequence of their poor ability to extract the statistical regularities in language input, due to their short-term or working memory limitations. Thus, although children with DLD are characterized by poor language skills, comorbid impairments in verbal working memory may underlie their impairments in verbal statistical or procedural learning tasks. It is possible, then, a verbal working memory deficit in DLD may be associated with poor performance on verbal- but not non-verbal statistical learning tasks.

3.1.1 Present study

In this study, school-aged children with and without DLD completed a SLL and VSL task, as well as a range of cognitive and linguistic measures. The SLL task was a word segmentation task (e.g., Saffran, Newport & Aslin, 1996), while the VSL task was a shape-triplet segmentation task (e.g., Siegelman et al., 2016). One goal of the study was to compare performance in cross-domain statistical learning tasks in groups differing primarily in relative language skills. Findings for impairments across both statistical learning tasks in the DLD group may point to either an impairment in processing structured information across domains, or an impairment in other domain-general processes associated with statistical learning. DLD deficits confined to the SLL task, on the other hand, would point to a domain-specific constraint in processing linguistic material for those with DLD, and would be consistent with prior reports (Lammertink et al., 2017).

A second aim of the study was to examine relationships between performance in statistical learning tasks to broader measures of language and other cognitive processes including working memory. Because verbal working memory has been shown to be impaired in children with DLD (Montgomery, 2002), with visuospatial working memory relatively spared, (Archibald & Gathercole, 2006a; Bavin, Wilson, Maruff & Sleeman, 2005), one could predict a domain-specific association between the poor verbal working memory abilities in those with DLD and poor SLL, and intact visuospatial working memory in this group being associated with unimpaired VSL. Relationships between domain-specific working memory skills and the SLL and VSL tasks would point to possible working memory involvement in statistical learning.

3.2 Method

3.2.1 Participants

Children were recruited from the developmental research pool at The University of Western Ontario (n = 11), and from those participating in a language and literacy intervention program in an elementary school in Southwestern Ontario (n = 16). Data from four children were excluded due to incomplete data collection, resulting in 23 total participants (11 females), M_{age} = 7.4 years, SD_{age} = 0.92 years, range: min_{age} = 5.25 years, max_{age} = 9.67 years. Nineteen children participated in follow-up testing six months after initial testing. All participants were native English speakers. Participants who came to the laboratory for testing received a $20 gift card as compensation for study participation. Participants recruited from the language and literacy intervention program received a language report from a speech language pathologist as compensation for study participation. Ethics approval for all study procedures and materials was obtained by the University of Western Ontario Non-Medical Research Ethics Board. Written informed consent was obtained from the parents of all study participants, and written and verbal assent was obtained from all study participants.

Twelve of the 23 children in the initial sample met the criteria for DLD; inclusionary criterion was one standard deviation (< 85) below average on the *Clinical Evaluation of Language Fundamentals* (CELF-IV; Semel, Wiig & Secord, 2003) Composite Language Score (CLS). The CLS includes scores from four core subtests from the CELF-IV, including *Concepts and Following Directions, Recalling Sentences, Formulating Sentences,* and, depending on the age of the children, *Word Knowledge* (< 9 years) or *Word Classes* (> or equal to 9 years). These tests are described in more detail below. Scores for the language measures and the other standardized measures are presented in Table 3.1.

Table 3.1 Standard scores for DLD and TD groups across the standardized language and cognitive measures.

Group	*n*	*Age (Years)*	*CELF*	*AWMA*	*MAVA Expressive*	*MAVA Receptive*	*WASI Block Design*	*WASI Matrix Reas.*
TD	12	7.73 (1.33)	100.00 (11.52)	96.42 (15.52)	101.67 (6.76)	94.50 (18.49)	52.82 (10.59)	54.82 (12.05)
DLD	11	7.33 (0.94)	66.18** (7.90)	83.32* (11.94)	89.09* (10.30)	85.00 (13.34)	43.11* (6.97)	41.11* (9.88)

Note. Where DLD < TD: * = $p < .05$; ** = $p < .001$.

3.2.2 Procedure

All participants completed three individual study sessions less than one hour long at either a university laboratory or in a quiet room in their school. The first 2 study sessions occurred approximately 1 week apart, and a third, follow-up session was completed 6 months later. At the first testing session, children completed standardized measures of language and working memory. At the second testing session, the children completed standardized measures of vocabulary and nonverbal reasoning, as well as the statistical language learning task. Nineteen children who returned for the follow-up session completed the VSL task. At all sessions, additional measures not reported here were completed. All tasks were administered by a trained research assistant.

3.2.3 Materials

3.2.3.1 Standardized measures

Oral language. As described as part of the inclusionary criteria for DLD, each child completed the four core subtests for the child's age for the Composite Language Score (CLS) from the CELF-IV (Semel et al., 2003). In the *Concepts and Following Directions* subtest, the child pointed to aspects of a pictorial display following a spoken instruction. For *Recalling Sentences*, the child repeated sentences of increasing length immediately after hearing them. For *Formulated Sentences*, a child was given a word or words and generated a spoken sentence in reference to a picture cue. If the child was between 5 and 8 years old, they completed the *Word Structure* subtest, in which the child completed a sentence with the grammatically correct form of a target word. If the child was between 9 and 12 years old, they completed the *Word* Classes subtest, in which the child heard four words, and chose two from that set which were related, and described their relation. Scores on the four subtests were aggregated to calculate the CLS, and converted to standard scores based on published norms.

Vocabulary. Children completed both the Receptive and Expressive subtests of the *Montgomery Assessment of Vocabulary Acquisition* (MAVA; Montgomery, 2008). For the *Receptive* subtests, a child was given a target word and had to select the corresponding picture from an array of four pictures. For the *Expressive* subtest, the child was shown a target picture and had to provide the correct name for the picture. For both MAVA subtests, raw scores were converted to standard scores based on published norms.

Working memory. Two subtests from the *Automated Working Memory Assessment* (AWMA; Alloway, 2007) were administered. The *Dot Matrix* subtest involved recalling the position of a series of dots in a 3x4 matrix. The *Counting Recall* subtest involved counting the number of circles in an array of shapes, and recalling the respective tallies of circles at the end of the list. Both tasks required the storage and processing of information in a span procedure such that up to six lists of one item, and then two items, and so on were completed until errors were made on four of the six lists, at which point the test was discontinued. Standardized scores were calculated based on local norms (Nadler & Archibald, 2014). These tests have been found to load on to separable working memory factors. Specifically, the *Dot Matrix* subtest is associated with visuospatial working memory, while the *Counting Recall* subtest is associated with verbal working memory and phonological storage (Archibald, 2013).

Nonverbal Reasoning. Children completed the *Block Design* and *Matrix Reasoning* subtests of the *Wechsler Abbreviated Intelligence Scale* (WASI, Wechsler, 2003). For the *Block Design* subtest, the child arranged blocks to match a model or a picture. For the *Matrix Reasoning* subtest, the child chose a picture to complete a 5-element pattern. For both WASI subtests, raw scores were converted to standard scores based on published norms.

3.2.3.2 Statistical learning tasks

Statistical language learning. Participants were exposed to a structured, unsegmented speech stream for 21-minutes, followed immediately by a two-alternative forced-choice (2AFC) test, asking participants to identify words from the artificial language. In the exposure phase, the artificial language was played over headphones at a comfortable listening volume. Participants were told that the experimenter was "going to play some sounds over their headphones", with no information provided about the artificial language. These deliberately vague instructions minimized the chance of participants trying to explicitly learn the language's structure during the experiment. During the artificial language exposure phase, participants were free to colour a colouring page.[1]

The artificial language was based on the stimuli described by Saffran, Newport and Aslin (1996). The language was composed of an inventory of 12 CV syllables, combined to create six trisyllabic "words": *patubi, tutibu, babupu, bupada, dutaba, pidadi*. Transitional probabilities of syllables ranged from 0.33 to 1.0 within-word, and from 0.1 to 0.2 across word boundaries, assuming an equal distribution of each word preceding and following all others. The artificial language was constructed from audio recordings of a female native-English speaker using a neutral vocal effort. Recordings of the speech stimuli were made in a double walled IAC sound booth with a pedestal microphone (AKG C 4000B) located approximately 30cm from the speaker's mouth and routed to a USBPre 2 pre-amplifier (Sound Devices) using SpectraPlus software (Pioneer Hill Software, 2008). Recordings were made of each of the 12 target syllables in the middle of a three-syllable sequence, within every co-articulation context required for the language. Eight repetitions of each sequence were recorded, and the token with the most neutral pitch contour and best sound quality was chosen and uploaded into Sound Forge Audio Studio (Sony Creative Software) editing software. Middle syllables from the recorded tokens were extracted by identifying the final offset of vowel oscillation in the

[1] In Saffran et al. (1997) and Evans et al. (2009), children were instructed to colour on a computer colouring program. Although this study aimed to closely replicate these original findings, experimental piloting suggested that the participants in the present study were more engaged with pencil-and-paper colouring compared to the computerized version.

previous syllable to the offset of vowel oscillation in the target syllable. These were then concatenated to create the final 21-minute stream of words. The stream consisted of 360 tokens of each word in random order, with no word presented twice in sequence. The language maintained a consistent speech rate (average 3.1 syllables/s) using a time stretch, and was normalized to a pitch of F0 = 196 Hz using the pitch shift in Sound Forge Audio Studio. There were no pauses between words; as such, the only cues to word boundaries were the lower transitional probabilities for between-word syllable pairs.

The 2AFC test followed immediately after the exposure phase, which was delivered by E-Prime 2.08 (Schneider, Eschman, & Zuccolotto, 2002). For each test item, participants heard a trained word from the artificial language paired with a non-word foil, separated by 500ms of silence. Presentation order of trained words and non-word foils were randomized across trials. During the test phase, participants were asked to select the word that "sounds more like something you heard in the language". Before the test phase began, participants completed four practice trials to ensure they understood the task. Then, the test phase began. Within the test phase, each non-word foil was paired exhaustively with each trained word, comprising 36 total test pairs. The test pairs were presented in a fixed random order. Behavioural accuracy on the task was calculated for each participant as the percent of correct identifications of trained words.

For the test phase stimuli, an additional six non-word foils were constructed from the same 12 CV syllables as the artificial language, but which were not included in the training set: *pubati, tapudi, dupitu, tipabu, bidata, batipi*. Non-word foils were created with the constraint that within-word transitional probabilities would be zero. Syllables were drawn from the same recording inventory as the artificial language stimuli, with appropriate co-articulation contexts. Note that fully new non-word foils were used with transitional probabilities of 0 based on previous language exposure (e.g., Evans et al., 2009), rather than tri-syllabic part-word foils consisting of two syllables from a trained item plus an incorrect syllable (e.g., Saffran, Newport, & Aslin 1996). Discrimination accuracy at test is generally higher when fully new non-word foils are used rather than part-words. The six non-word foils were paired exhaustively with the six trained words, creating 36 total test pairs.

Visual statistical learning. Similar to the statistical language learning task, the VSL task involved an exposure phase and a test phase, and was adapted from the task designed by Siegelman, et al. (2016) to be used with a developmental population. Participants were seated in front of a laptop computer, and were told that they would see some shapes one at a time presented on the computer. As with the SLL task, deliberately vague instructions were provided for the VSL task. While the shape stream was presented on the computer, the experimenter played the instrumental soundtrack from the film "Inside Out" (Inside Out Original Motion Picture Soundtrack, 2015) over a speaker. Playing music concurrent with the VSL exposure phase was done to closely approximate the SLL exposure phase by including a non-demanding secondary task well-suited to school-aged children. The VSL task was administered using E-Prime 2.08 (Schneider et al., 2002)

For the exposure phase stimuli, 10 complex black shapes were concatenated into five triplets (see Appendix 1). The transitional probabilities between adjacent shapes within a triplet ranged from 0.33 to 1.0. The five triplets were concatenated into a random stream, with the restriction of no immediate repetitions of triplets. Each triplet appeared 24 times in the familiarization phase. Shapes were presented for 800ms, with a 200ms inter-stimulus-interval between shapes. Thus, the stream was 6 minutes in length. Note that the present task was modified from the original paradigm to be used with school-aged children by including fewer shapes in the training set (10 as opposed to 16), and fewer triplets (six, as opposed to eight).

Following familiarization, participants completed a 35-item test. All test items are detailed in Appendix 2. The test phase was shortened from the 42-item test used in Siegelman et al. (2016) in order to accommodate having fewer training triplets and fewer shapes in the training set than the original paradigm. For the test phase stimuli, an additional eight foil triplets were constructed from the same 10 shapes used in the training set. The transitional probabilities of the foil items ranged between 0 to 0.5 and differed in their position violations That is, a violation occurred for either the onset, medial, or final shape of the sequence across items.

The test phase consisted of pattern recognition and pattern completion items. Pattern recognition items included two- and four-alternative forced choice questions for both triplet and pair sequences from the training stream. Answers were clearly numbered on the computer screen, and participants were asked to select the corresponding number for their answer while an experimenter coded their responses. Pattern completion items included questions for both triplet and pair sequences from the training stream. For these items, participants selected the missing shape to complete a sequence from the training set from a selection of three possible shapes. As with the pattern recognition items, participants indicated their answer and the experimenter coded their responses. The corresponding instructions for each question were presented on the computer screen during the test phase. For the pattern recognition items, the instruction read “Please choose the pattern you are most familiar with as a whole”. For the pattern completion items, the instruction read “Please choose the shape that best completes the pattern”. A visualization of a selection of test items is provided in Figure 3.1.

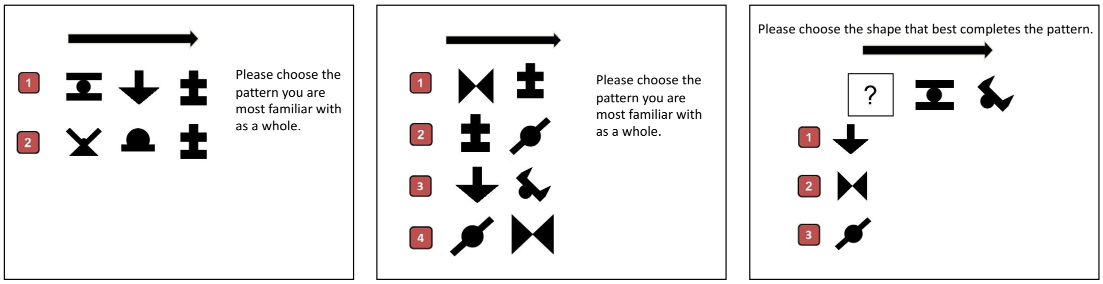


Figure 3. 1. Sample questions from the visual statistical learning (VSL) task test phase. From left to right, the panels depict a triplet-2AFC, a pair-4AFC, and a triplet pattern completion question.

The test began with the pattern recognition items, which consisted of 10 triplet-2AFC items, followed by five triplet-4AFC items, five pair-2AFC items, and five pair-4AFC items. Finally, there were five pattern completion items for triplet sequences, and five pattern completion items for pair sequences. The order of presentation of items within each question type was randomized across participants. The correct response (either item “1” or “2” for 2AFC questions, or “1”, “2”, “3”, or “4” for 4AFC questions) was counterbalanced amongst items. Each trained (target) triplet appeared as a target in the pattern completion items for both triplet and pair completions. Note that the presentation of trained triplets and foil items was not fully exhaustive, as it was in the statistical language learning test phase. Following Siegelman et al. (2016), test items such as these both maximized the number of test items and minimized the number of repeated targets and foils in order to mitigate any learning effects within the test phase.

3.3 Results

3.3.1 Statistical language learning

First, a one-sample *t*-test was performed to compare performance on the two-alternative forced-choice (2AFC) test against chance. Individual scores are presented in

Figure 3.2. Chance on the 2AFC was defined as achieving 18 (50%) correct out of the 36 total test items. Bootstrapping with 10,000 samples was used to validate the results of the t-test. Performance on the 2AFC task was above chance, $M = 53.65\%$, $SD = 7.65\%$, $t\ (22) = 2.290$, $p = 0.032$, $d = 0.477$. The data were also examined by estimating a Bayes factor using Bayesian Information Criteria (Wagenmakers, 2007), comparing the fit of the data under the null hypothesis and the alternative hypothesis. An estimated Bayes factor (alternative/null) suggested that the data were 2.18:1 in favor of the alternative hypothesis, or rather, 2.18 times more likely to occur under a model including an effect of statistical language learning, rather than a model without it. This represents positive (Raferty, 1995) or substantial (Jeffreys, 1961) evidence that scores on the 2AFC were above chance. Because the effect was generally small, performance on individual test items was also examined. Scores on each individual test item, averaged across participants, was compared to chance (50%) using a t-test. As with the previous analyses, bootstrapping with 10,000 samples was used for these tests. Performance on only two test items was significantly above chance: *tipabu* versus *pidadi*, $M = 70.00\%$, $SD = 46.5\%$, $t\ (22) = 2.554$, $p = .018$, $d = 0.430$, and *tipabu* versus *babupu,* $M = 78.00\%$, $SD = 42.4\%$, $t\ (22) = 3.407$, $p = .002$, $d = 0.660$. Above-chance performance was not achieved on any of the other test items, $t\ (22) < 1.499$, $p > .05$, $d < 0.354$, *all cases*.

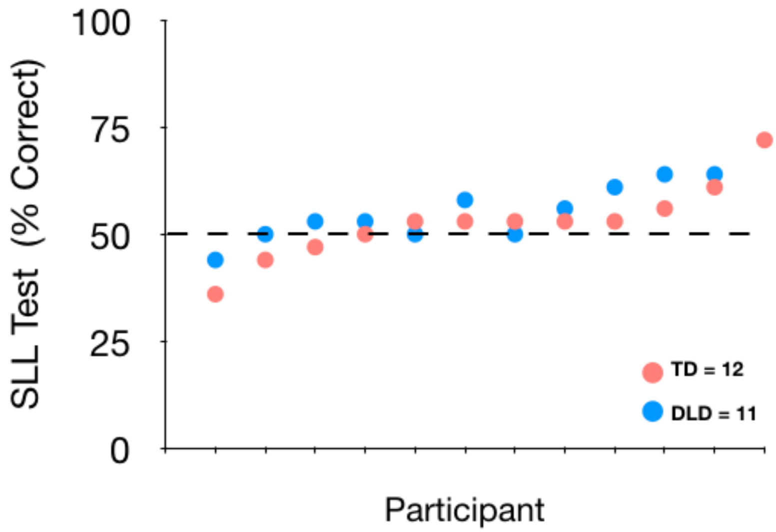


Figure 3.2 Performance on the 2AFC test of statistical language learning. The horizontal dashed line represents chance responding (50% correct).

In order to examine relationships between statistical language learning and vocabulary, language, working memory, and nonverbal reasoning, bivariate correlations were conducted (*see* Appendix 3). No significant correlations were observed between statistical language learning and age, nor any of the standardized measures, $r < 0.297$, $p > .05$, $BF_{10} < 0.715:1$, *all cases*. There were a number of significant correlations amongst the standardized measures, but given that these correlations were not pertinent to the present study, they were not considered further.

Taken together, these results showed generally weak performance on the statistical language learning task, with above-chance performance on only two of the test items. Furthermore, given these data, there was insufficient evidence to support an association between scores on the statistical language learning task and age, nor any of the standardized cognitive measures.

3.3.1.1 Comparing children with and without DLD on statistical language learning

The DLD (M = 54.82%; SD = 6.42%) and TD (M = 52.83%; SD = 8.77%) groups did not differ on the 2AFC statistical language learning measure, $t\ (21) = -0.692$, p = .497[2], $d = 0.002$. The DLD group was above chance on the SLL measure, $t\ (10) = 2.491$, $p = .032$, $d = 0.750$, while the TD group was not above chance, $t\ (11) = 1.020$, $p = .330$, $d = 0.323$.

Similar to the sample-level analyses, the group-level data were also examined by estimating a Bayes factor. For the comparison of the DLD and TD group, the estimated Bayes factor (null/alternative) suggested that the data were 2.219:1 in favor of the null hypothesis, or rather, 0.45 times more likely to occur under a model not including an effect group, rather than a model with it. Thus, there was little evidence to suggest that the groups differed on the SLL task.

Taken together, these results clearly demonstrated that the DLD and TD groups did not differ on the 2AFC task assessing statistical word segmentation. Additionally, there was some evidence that the DLD group was numerically above-chance on the SLL task. However, given the small sample size for this analysis, the results should be interpreted with caution.

3.3.2 Visual statistical learning

Scores across participants on the VSL task are presented in Figure 3.3. Because the number of response options varied amongst question types, chance could not be

---

[2] Given that the sample sizes under consideration were rather small (DLD: n = 11; TD n = 12), a power analysis was conducted using G-Power (*version* 3.1.9.3, 2009) for the t-test comparing statistical language learning scores between groups. The power for the t-test was 0.7546, which approached the standard of adequacy at 0.80. Thus, it was warranted to assume that there was a failure to reject the null hypothesis.

defined as 50% correct on the test as a whole. According to the binomial distribution, the different probabilities of correct responses on different items were aggregated to determine that chance-level performance was 13.33 correct trials (e.g., Siegelman et al., 2016). Thus, above-chance performance was defined as scoring above 13.33 correct responses, or 38.09% correct. At the group level, performance did not differ from chance, $M = 39.09\%$, $SD = 8.78\%$, $t\ (20) = 0.740$, $p = .468$, $d = 0.114$. An estimated Bayes factor (null/alternative) suggested that the data were 3.44:1 in favor of the null hypothesis. Similar to the analyses of item-level performance on the statistical language learning task, scores on each individual test item were compared to chance. The appropriate chance-level criterion was selected for individual questions. Results showed that performance was not above chance for any question, $t\ (21) < 2.046$, $p > .05$, $d < 0.558$, *all cases*. Thus, these results clearly demonstrated that group-level performance across the VSL test was not above chance.

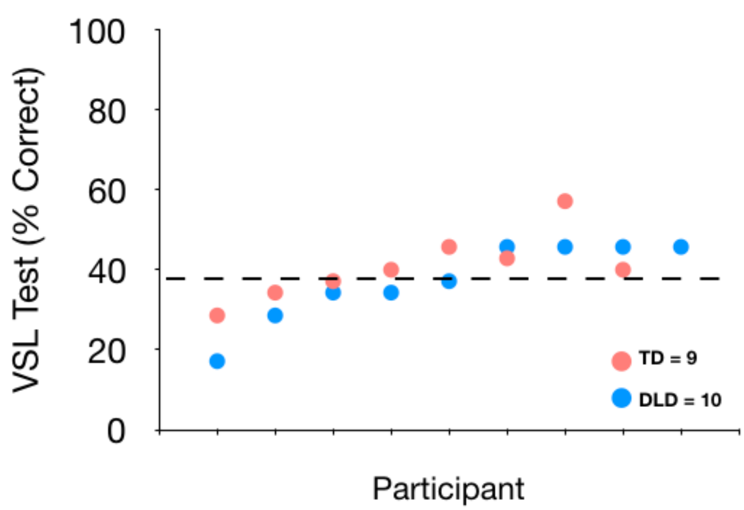


Figure 3.3 Individual performance on the test of visual statistical learning (VSL). The horizontal dashed line represents chance responding (38.09% correct).

As with the analyses of the SLL data, a series of correlations were completed comparing average scores on VSL task against age, and the standardized measures, as well as SLL scores (*see* Appendix 4). Performance on the VSL task did not correlate with SLL scores, age, nor any of the cognitive measures, $r < .247$, $p > .05$, $BF_{10} < 1.085$ *all cases*. There were a number of significant correlations amongst the standardized measures, but they were not considered further for the present study.

Taken together, performance on the VSL task was not above chance. Furthermore, given these data, there was insufficient evidence to support an association between performance on the VSL task and performance on the SLL task, nor any of the cognitive abilities assessed in the present study.

3.3.2.1 Comparing children with and without DLD on visual statistical learning

As the sample included in these analyses is not identical to the sample used in comparing statistical language learning scores, descriptive statistics for age, cognitive measures, and VSL for children with DLD and TD children are presented again here, in Table 3.2. Children with DLD scored significantly lower than their TD peers on the measures of language, $t\,(17) = 6.690$, $p < .001$, expressive vocabulary, $t\,(17) = 2.857$, $p = .011$, working memory, $t\,(17) = 2.446$, $p = .026$, and *Block Design*: $t\,(17) = 2.707$, $p = .017$. The groups did not differ on receptive vocabulary, *Matrix Reasoning*, or age, $t\,(17) < 2.083$, $p > .05$, *all cases*.

Table 3.2 Scores for DLD and TD groups in the follow-up sample across standardized and cognitive measures.

Group	*n*	*Age (Years)*	*CELF*	*AWMA*	*MAVA Expressive*	*MAVA Receptive*	*WASI Block Design*	*WASI Matrix Reas.*
TD	9	7.73 (1.33)	99.89 (12.62)	95.22 (13.92)	101.22 (7.42)	88.89 (16.37)	52.82 (10.59)	54.82 (12.05)
DLD	10	7.33 (0.94)	66.18** (7.90)	83.32* (11.94)	89.09* (10.30)	85.00 (13.34)	43.11* (6.97)	41.11 (9.88)

Note. Where DLD < TD: * = $p < .05$; ** = $p < .001$.

Performance on the VSL task did not differ between the TD (*M* = 41.30%, *SD* = 8.1%) and DLD (*M* = 37.1%, *SD* = 9.3%) groups, t (17) = 1.175, p = .256[3], d = 0.482. Also, performance for neither group was statistically above chance, t < 1.178, p > .272, d < .0.396, all *cases*. Following the approach used to examine the group comparison on the SLL measure, the VSL task performance data were also compared between TD and DLD groups using a Bayes factor. The estimated Bayes factor (null/alternative) suggested that the data were 1.72:1 in favour of the null hypothesis. Thus, there was no evidence to suggest that the TD and DLD groups differed in performance on the VSL task. Additionally, the Bayes factors estimating whether the TD or DLD groups were above chance on the VSL task provided no evidence in favour of the alternative hypotheses (alternative/null factor: < 0.556:1, *all cases*).

Overall, then, these results clearly demonstrated that performance on the VSL task did not differ for the DLD and TD groups, and that neither group scored above chance.

3.4 Discussion

In the present study, children with developmental language disorder (DLD) were compared to typically developing (TD) children in their performance on a statistical language learning (SLL) and visual statistical learning (VSL) task. The relationship between performance on these two statistical learning tasks and other cognitive and linguistic measures was also examined. Analysis of performance on the SLL task revealed that, across the sample, performance was marginally above chance. Comparing performance between DLD and TD groups, it was found that the two groups did not

---

[3] Given that the sample sizes under consideration were again rather small (DLD: n = 10; TD n = 9), a power analysis was conducted using G-Power (*version* 3.1.9.3, 2009) for the t-test comparing VSL scores between groups. The power for the t-test was 0.7569, which approached the standard of adequacy at 0.80. Thus, it was warranted to assume that there was a failure to reject the null hypothesis.

differ. However, the DLD group performed above chance, while the TD group did not. On the VSL task, performance across the sample was not above chance, and the TD and DLD groups did not differ. Additionally, performance on the two statistical learning tasks was not related, nor was there evidence to conclude that there was a relationship between performance on either statistical learning task related to any of the cognitive or linguistic measures. These results were a failure to replicate previous work (Evans et al., 2009) which demonstrated a group difference on statistical word segmentation between TD and DLD groups and an association between SLL and vocabulary. This work also failed to show that school-aged children could learn the distributional regularities within an adapted VSL task.

The results from the present study, with only the DLD group being above-chance on the SLL task with neither group above-chance on the VSL task, failed to clearly address the main research question investigating the domain-specificity or domain-generality of statistical learning within a population with a disproportional language impairment. The finding of no group difference on the SLL task between the TD and DLD groups was surprising, given that a statistical or procedural learning deficits have repeatedly been demonstrated in the research (e.g., Lammertink et al., 2017; Lum et al., 2014; Obeid et al., 2016). The above-chance performance on the SLL task for the DLD group also runs contrary to these previous findings. Nevertheless, both a power analysis and a Bayes' factor calculated for the present effect confirmed the findings. A likely conclusion is that the above-chance performance for the DLD group is a weak and precarious effect, and further investigation would be necessary to demonstrate successful statistical word segmentation for this group. The results for the SLL task from the present study should also be interpreted with caution given the relatively small effect and small sample size. Although the accuracy rate on the 2AFC of 53% is consistent with prior research (e.g., Evans et al., 2009), it is difficult to conclude whether this rate can be taken as convincing evidence of learning. The lack of a learning effect on the VSL task for either group is also difficult to interpret, and may reflect measurement error or an unsuitability of this task for a developmental sample. In order to make any conclusions regarding a domain-specific or domain-general impairment on statistical learning in the language impaired group, above-chance performance for the TD group on either statistical learning task would be necessary.

Although it is difficult to claim from the present results that the sample "learned" the regularities in either the verbal or visual domain, some participants were nevertheless above chance. Thus, it was interesting to speculate as to whether performance on either task was related. What was concluded from the present results was that performance on the SLL and VSL tasks was unrelated. The lack of a correlation among performance outcomes on these tasks may be related to low learning overall, but, nevertheless, this finding has some relevance to prior research. Directly comparing performance on a SLL and VSL task, Siegelman and Frost (2015) reported that the correlation between these two tasks was virtually zero. Examining the developmental trajectory of verbal and visual statistical learning, Raviv and Arnon (2017) showed that while performance on a visual SL task improved linearly with age (*also see* Arciuli & Simpson, 2011), auditory SL did not show any improvements with age. The authors reasoned that there are different developmental trajectories across these two tasks and, importantly, that there are clear domain-specific constraints on SL. It is then reasonable to conclude that the lack of a correlation between visual and auditory SL in the present study stems from different constraints in processing regularities in either domain. However, this conclusion should be further evaluated with more sensitive measures where learning across both tasks can clearly be demonstrated.

One alternative to explain the lack of a correlation between performance on the SLL and VSL tasks, and above-chance performance on only the SLL task, is that the SLL task was subject to linguistic entrenchment (Siegelman, Bogaerts, Elazar, Arciuli, & Frost, 2018). The linguistic entrenchment hypothesis suggests that learning in a statistical learning task is constrained by prior knowledge. Specifically, learners' entrenched expectations about co-occurrences from their native language impact what they learn from novel auditory input. Learning in the VSL task, however, is not shaped by prior knowledge because participants should, presumably, have no prior knowledge about the nonsense shape sequences used in this task. Investigating the linguistic entrenchment hypothesis, Siegelman and colleagues (2018) found that auditory-verbal statistical learning tasks display item-specific learning effects, and that performance on auditory-verbal statistical learning tasks is related to participants' native-language knowledge. In the present study, although the nonsense words do not exist in the participants' lexicons, they would be consistent with existing linguistic knowledge. For instance, the syllables used in the present study follow the phonotactic constraints within English. So, it is possible that prior knowledge of some aspects the linguistic stimuli affected performance on the SLL task, while the VSL task was unaffected due to a lack of prior experience. This may explain why performance on the SLL task was above chance, while performance on the VSL task was not. Yet, an entrenchment explanation remains unlikely in the present study, where above-chance performance on the SLL task was seen only for the DLD group. Children with DLD typically present with low vocabularies (Laws & Bishop, 2003), and underspecified phonetic representations (Edwards & Lahey, 1996).

These reports are thus difficult to reconcile with their performance on the SLL being affected by linguistic entrenchment.

Perhaps the greatest implication that can be taken from these inconclusive results is that there are improvements to be made to the design of statistical learning paradigms, particularly for use in developmental samples. Batterink, Reber, Neville, and Paller (2015) demonstrated that traditional statistical learning assessments based on familiarity judgments, including the 2AFC, reflect explicit memory, and should not be taken as measures of implicit learning. Their results also demonstrated that indirect measures capture implicit learning more effectively. Because statistical learning paradigms are primarily concerned with uncovering implicit learning effects, a better measurement approach is warranted. To address this, a number of implicit measures of statistical learning have been described in the literature. Batterink and colleauges (2015), developed a measure that involves measuring reaction times (RTs) to syllables within a structured, unsegmented speech stream, similar to the one used in the present paradigm. Findings across related studies (e.g., Batterink, Reber, & Paller, 2015; Batterink & Paller, 2017) revealed that participants were faster at responding to learned stimuli, which was reflective of learning the distributional regularities within the trained artificial language (*also see* Misyak, Christiansen, & Tomblin, 2010). A second possible approach is to measure event-related potentials (ERPs) during a statistical learning paradigm. Previous research has shown that following language exposure, participants' ERP responses are reflective of successful word segmentation (e.g., Cunillera, Toro, & Sebastián-Gallés, 2006; Cunillera et al., 2009; Sanders, Newport, Neville, 2002). A final approach is to examine participants' brain responses to structured stimuli by measuring fMRI during statistical learning. fMRI approaches have largely been adopted to measure VSL (Turk-

Browne, Scholl, Chun, & Johnson, 2009; Schapiro, Gregory, Landau, McCloskey, & Turk-Browne, 2014) although some studies have also used fMRI to measure SLL (e.g. Cunillera et al., 2009; Scott-van Zeeland, McNealy, Wang, Sigman, Brookheimer, & Dapretto, 2010; Plante, Patterson, Dailey, Kyle, & Fridriksson, 2014;). Given the promise of this research, future research should aim to examine implicit statistical learning in a developmental sample using an indirect technique.

Another possibility as to why the results of the SLL paradigm from Evans et al. (2009) were not replicated in the present study is that the effect for this paradigm is meager, or truly not replicable. There are no reported findings showing the absence of a learning effect in TD school-aged children on the SLL paradigm from Saffran et al. (1997). However, a number of studies have tried to extend the findings from the original Saffran, Aslin, and Newport (1996) paradigm in infants, and have failed to show a learning effect. Both Johnson and Jusczyk (2003) and Lew-Williams and Saffran (2012) showed that infants could only segment a speech steam when words were of uniform length, and failed to segment the speech stream when words were of varying length. Johnson and Jusczyk (2003) suggested that the simplicity of the language in Saffran, Aslin, and Newport (1996) is what drove the effect, and that the statistical learning phenomenon exhibited in the original paradigm fails to scale up to natural language where, indeed, words are of varying length. Additionally, Graf-Estes and Lew-Williams (2015) demonstrated that infants failed to segment words when different voices were used in training and at test, providing further evidence that the original Saffran, Aslin, and Newport (1996) paradigm is difficult to replicate when complexity is added to the input stream. Thus, there is some speculation to be raised about the replicability of the word segmentation effect originally proposed by Saffran, Aslin, and Newport (1996).

What is possible is that additional cues may aid in successful segmentation. Some have argued that hearing words in isolation may help bootstrap infants' speech segmentation (Brent, 1999; Brent & Siskind, 2001; Pinker, 1984). Others have argued that highly frequent words, such as names and function words, may further facilitate speech segmentation (Bortfeld, Morgan, Golinkoff, & Rathburn, 2005; Shi, Cutler, Werker & Cruickshank, 2006). Consistent with this, research on a large-scale language has shown that adults can successfully segment words from a speech stream constructed with a Zipfian distribution (Kurumada, Meylan & Frank, 2010), albeit with 10 hours of language exposure. Still, others have argued that prosody is essential to early word segmentation (Johnson, 2008; Johnson & Jusczyk, 2001; Jusczyk, 1997; Johnson & Seidl, 2008, 2009; Mehler, Nespor, & Shukla, 2006), or may contribute to word segmentation after the onset of transitional probability tracking (Thiessen & Saffran, 2007). Although much of the work reviewed here has focused on infant research, it is nevertheless informative for exploring some of the potential shortcomings of the statistical word segmentation paradigm used in the present study. What can be said for these findings is that tracking transitional probabilities between syllables (or shapes) is likely not the only way in which an individual segments the regularities in the environment. Indeed, no one has ever claimed that transitional probabilities are the sole source of information used in successful segmentation. What may be the case is that the conjunction of segmentation cues, including cues such as prosody and transitional probabilities in SLL paradigms, would facilitate segmentation in the present paradigm.

3.4.1 Conclusions

In order to examine the domain-specificity or domain-generality of statistical learning, the present study investigated statistical learning across both verbal and visual statistical learning tasks in children with DLD. For the SLL task, performance for children with DLD was above chance, while performance for TD children was not. For the VSL task, performance for neither the DLD and TD groups was above chance. The associations between statistical learning and an array of cognitive and linguistic measures was also examined, and these data did not support an association between statistical learning and any of these broader measures. Additionally, performance on the two statistical learning tasks was not associated. Overall, these results do not provide sufficient evidence to conclude that statistical learning is mediated by domain-general or domain-specific processes. However, they do highlight the necessity for developing more sensitive measures of implicit statistical learning that can accurately capture learning in these tasks.

3.5 References

Alloway, T. P. (2007). Automated Working Memory Assessment. London: Pearson Assessment.

Altmann, G., Dienes, Z., & Goode, A. (1995). Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 899.

Archibald, L. M. (2013). The language, working memory, and other cognitive demands of verbal tasks. *Topics in Language Disorders*, *33*(3), 190-207.

Archibald, L. M., & Gathercole, S. E. (2006a). Visuospatial immediate memory in specific language impairment. *Journal of Speech, Language, and Hearing Research*, *49*(2), 265-277.

Archibald, L. M., & Gathercole, S. E. (2006b). Short-term and working memory in specific language impairment. *International Journal of Language & Communication Disorders*, *41*(6), 675-693.

Archibald, L. M., & Joanisse, M. F. (2009). On the sensitivity and specificity of non-word repetition and sentence recall to language and memory impairments in children. *Journal of Speech, Language, and Hearing Research*, *52*(4), 899-914.

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

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9,* 321-324.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*(1), 158.

Batterink, L. J., & Paller, K. A. (2017). Online neural monitoring of statistical learning. *Cortex*, *90*, 31-45.

Batterink, L. J., Reber, P. J., & Paller, K. A. (2015). Functional differences between statistical learning with and without explicit training. *Learning & Memory*, *22*(11), 544-556.

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language*, *83*, 62-78.

Bavin, E. L., Wilson, P. H., Maruff, P., & Sleeman, F. (2005). Spatio-visual memory of children with specific language impairment: evidence for generalized processing problems. *International Journal of Language & Communication Disorders*, *40*(3), 319-332.

Bishop, D. V. (1979). Comprehension in developmental language disorders. *Developmental Medicine & Child Neurology*, *21*(2), 225-238.

Bishop, D. V. (1982). Comprehension of spoken, written and signed sentences in childhood language disorders. *Journal of Child Psychology and Psychiatry*, *23*(1), 1-20.

Bishop, D. V. (1997). Cognitive neuropsychology and developmental disorders: Uncomfortable bedfellows. *The Quarterly Journal of Experimental Psychology Section A*, *50*(4), 899-923.

Bishop, D. V., North, T., & Donlan, C. H. R. I. S. (1996). Non-word repetition as a behavioural marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry*, *37*(4), 391-403.

Bishop, D. V., Snowling, M. J., Thompson, P. A., Greenhalgh, T., Catalise-2 Consortium, Adams, C., ... & Boyle, C. (2017). Phase 2 of CATALISE: A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry*, *58*(10), 1068-1080.

Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychological science*, *16*(4), 298-304.

Brent, M. R. (1999). Speech segmentation and word discovery: A computational perspective. *Trends in Cognitive Sciences*, *3*(8), 294-301.

Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, *81*(2), B33-B44.

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

Choudhury, N., & Benasich, A. A. (2003). A family aggregation study: the influence of family history and other risk factors on language development. *Journal of Speech, Language, and Hearing Research*, *46*(2), 261-272.

Christiansen, M. H. (2018). Implicit statistical learning: a tale of two literatures. *Topics in Cognitive Science*.

Coady, J. A., & Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders*, *43*(1), 1-40.

Conti-Ramsden, G. (2003). Processing and linguistic markers in young children with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research*, *46*(5), 1029-1037.

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry*, *42*(6), 741-748.

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

Cunillera, T., Càmara, E., Toro, J. M., Marco-Pallares, J., Sebastián-Galles, N., Ortiz, H., ... & Rodríguez-Fornells, A. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *Neuroimage*, *48*(3), 541-553.

Cunillera, T., Toro, J. M., Sebastián-Gallés, N., & Rodríguez-Fornells, A. (2006). The effects of stress and statistical cues on continuous speech segmentation: an event-related brain potential study. *Brain Research*, *1123*(1), 168-178.

Dollaghan, C., & Campbell, T. F. (1998). Non-word repetition and child language impairment. *Journal of Speech, Language, and Hearing Research*, *41*(5), 1136-1146.

Edwards, J., & Lahey, M. (1998). Non-word repetitions of children with specific language impairment: Exploration of some explanations for their inaccuracies. *Applied Psycholinguistics*, *19*(2), 279-309.

Estes, K. G., & Lew-Williams, C. (2015). Listening through voices: Infant statistical word segmentation across multiple speakers. *Developmental Psychology*, *51*(11), 1517.

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

Fedorenko, E., & Thompson-Schill, S. L. (2014). Reworking the language network. *Trends in Cognitive Cciences*, *18*(3), 120-126.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences*, *99*(24), 15822-15826.

Flax, J. F., Realpe-Bonilla, T., Hirsch, L. S., Brzustowicz, L. M., Bartlett, C. W., & Tallal, P. (2003). Specific language impairment in families: Evidence for co-occurrence with reading impairments. *Journal of Speech, Language, and Hearing Research*, *46*(3), 530-543.

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

Gathercole, S. E. (2006). Non-word repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*(4), 513-543.

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

Hamrick, P., Lum, J. A., & Ullman, M. T. (2018). Child first language and adult second language are both tied to general-purpose learning systems. *Proceedings of the National Academy of Sciences*, *115*(7), 1487-1492.

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*(3), B53-B64.

Hsu, H. J., & Bishop, D. V. (2010). Grammatical difficulties in children with specific language impairment: Is learning deficient?. *Human Development*, *53*(5), 264-277.

Jeffreys, H. (1961). Theory of probability (3rd ed.). New York, NY: Oxford University Press.

Johnson, E. K., & Jusczyk, P. W. (2003). Exploring statistical learning by 8-month-olds: The role of complexity and variation. *Jusczyk Lab final report*, 141-148.

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, E. K., & Seidl, A. (2008). Clause segmentation by 6-month-old infants: A crosslinguistic perspective. *Infancy*, *13*(5), 440-455.

Johnson, E. K., & Seidl, A. (2008). Clause segmentation by 6-month-old infants: A crosslinguistic perspective. *Infancy*, *13*(5), 440-455.

Johnson, E. K., & Seidl, A. H. (2009). At 11 months, prosody still outranks statistics. *Developmental science*, *12*(1), 131-141.

Jusczyk, P. W. (1997). Finding and remembering words: Some beginnings by English-learning infants. *Current Directions in Psychological Science*, *6*(6), 170-174.

Kurumada, C., Meylan, S. C., & Frank, M. C. (2013). Zipfian frequency distributions facilitate word segmentation in context. *Cognition*, *127*(3), 439-453.

Lammertink, I., Boersma, P., Wijnen, F., & Rispens, J. (2017). Statistical Learning in Specific Language Impairment: A Meta-Analysis. *Journal of Speech, Language, and Hearing Research*, *60*(12), 3474-3486.

Leonard, L. B. (1998). *Children with specific language impairment*. MIT press.

Leonard, L. B., Miller, C. A., Grela, B., Holland, A. L., Gerber, E., & Petucci, M. (2000). Production operations contribute to the grammatical morpheme limitations of children with specific language impairment. *Journal of Memory and Language*, *43*(2), 362-378.

Lew-Williams, C., & Saffran, J. R. (2012). All words are not created equal: Expectations about word length guide infant statistical learning. *Cognition*, *122*(2), 241-246.

Lopez-Barroso, D., de Diego-Balaguer, R., Cunillera, T., Camara, E., Münte, T. F., & Rodríguez-Fornells, A. (2011). Language learning under working memory constraints correlates with microstructural differences in the ventral language pathway. *Cerebral Cortex*, *21*(12), 2742-2750.

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

Mainela-Arnold, E., & Evans, J. L. (2014). Do statistical segmentation abilities predict lexical-phonological and lexical-semantic abilities in children with and without SLI?. *Journal of child language*, *41*(2), 327-351.

Majerus, S., Van der Linden, M., Mulder, L., Meulemans, T., & Peters, F. (2004). Verbal short-term memory reflects the sublexical organization of the phonological language network: Evidence from an incidental phonotactic learning paradigm. *Journal of Memory and Language*, *51*(2), 297-306.

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

Marton, K., Eichorn, N., Campanelli, L., & Zakarias, L. (2016). Working memory and interference control in children with specific language impairment. *Language and Linguistics Compass*, *10*(5), 211-224.

Mayor-Dubois, C., Zesiger, P., Van der Linden, M., & Roulet-Perez, E. (2014). Nondeclarative learning in children with specific language impairment: predicting regularities in the visuomotor, phonological, and cognitive domains. *Child Neuropsychology*, *20*(1), 14-22.

McGregor, K. K., Newman, R. M., Reilly, R. M., & Capone, N. C. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *45*(5), 998-1014.

McNealy, K., Mazziotta, J. C., & Dapretto, M. (2006). Cracking the language code: neural mechanisms underlying speech parsing. *Journal of Neuroscience*, *26*(29), 7629-7639.

Mehler, J., Nespor, M., Shukla, M., & Peña, M. (2006). Why is language unique to humans?. In *Novartis Foundation Symposium*, *270*, 251. Chichester; NY; John Wiley.

Meulemans, T., & Van der Linden, M. (2003). Implicit learning of complex information in amnesia. *Brain and Cognition*, *52*(2), 250-257.

Miller, C. A., Leonard, L. B., & Finneran, D. (2008). Grammaticality judgements in adolescents with and without language impairment. *International Journal of Language & Communication Disorders*, *43*(3), 346-360.

Milne, A. E., Wilson, B., & Christiansen, M. H. (2018). Structured sequence learning across sensory modalities in humans and nonhuman primates. *Current Opinion in Behavioral Sciences*, *21*, 39-48.
Misyak, J. B., Christiansen, M. H., & Tomblin, J. B. (2010). On-line individual differences in statistical learning predict language processing. *Frontiers in Psychology*, *1*, 31.
Montgomery, J. K. (2008). *Montgomery assessment of vocabulary acquisition*. Super Duper® Publications.
Montgomery, J. W. (1995). Sentence comprehension in children with specific language impairment: The role of phonological working memory. *Journal of Speech, Language, and Hearing Research*, *38*(1), 187-199.
Montgomery, J. W. (2002). Understanding the language difficulties of children with specific language impairments: Does verbal working memory matter?. *American Journal of Speech-Language Pathology*, *11*(1), 77-91.
Montgomery, J. W. (2003). Working memory and comprehension in children with specific language impairment: What we know so far. *Journal of Communication Disorders*, *36*(3), 221-231.
Nadler, R. T., & Archibald, L. (2014). The assessment of verbal and visuospatial working memory with school age Canadian children. *Revue Canadienne d'Orthophonie et d'Audiologie. Vol*, *38*(3).
Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*(1), 1-32.
Noonan, N. B., Redmond, S. M., & Archibald, L. M. (2014). Contributions of children's linguistic and working memory proficiencies to their judgments of grammaticality. *Journal of Speech, Language, and Hearing Research*, *57*(3), 979-989.
Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., ... & Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: evidence from a population study. *Journal of Child Psychology and Psychiatry*, *57*(11), 1247-1257.
Obeid, R., Brooks, P. J., Powers, K. L., Gillespie-Lynch, K., & Lum, J. A. (2016). Statistical learning in specific language impairment and autism spectrum disorder: A meta-analysis. *Frontiers in Psychology*, *7*, 1245.
Palmer, S. D., & Mattys, S. L. (2016). Speech segmentation by statistical learning is supported by domain-general processes within working memory. *The Quarterly Journal of Experimental Psychology, 69*(12), 2390-2401.
Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233-238.
Pinker, S. (1984). *Language Learning and Language Development.* Cambridge, Mass.: Harvard University Press.
Pioneer Hill Software. (2008). *Spectraplus* (Version 5.0.26.0). Poulsbo, Washington: Pioneer Hill.
Plante, E., Patterson, D., Dailey, N. S., Kyle, R. A., & Fridriksson, J. (2014). Dynamic changes in network activations characterize early learning of a natural language. *Neuropsychologia*, *62*, 77-86.
Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, 111-163.

Raviv, L., & Arnon, I. (2018). The developmental trajectory of children's auditory and visual statistical learning abilities: modality-based differences in the effect of age. *Developmental Science*, *21*(4), e12593.

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

Redington, M., & Chater, N. (1996). Transfer in artificial grammar learning: A reevaluation. *Journal of Experimental Psychology: General*, *125*(2), 123.

Rice, M. L., Wexler, K., & Redmond, S. M. (1999). Grammaticality judgments of an extended optional infinitive grammar: Evidence from English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *42*(4), 943-961.

Saffran, J. R. (2001). Words in a sea of sounds: The output of infant statistical learning. *Cognition*, *81*(2), 149-169.

Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current directions in psychological science*, *12*(4), 110-114.

Saffran, J. R., & Thiessen, E. D. (2007). Domain-general learning capacities. *Blackwell Handbook of Language Development*, 68-86.

Saffran, J. R., & Wilson, D. P. (2003). From syllables to syntax: multilevel statistical learning by 12-month-old infants. *Infancy*, *4*(2), 273-284.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 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*(1), 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., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, *8*(2), 101-105.

Sanders, L. D., Newport, E. L., & Neville, H. J. (2002). Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. *Nature Neuroscience*, *5*(7), 700.

Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M., & Turk-Browne, N. B. (2014). The necessity of the medial temporal lobe for statistical learning. *Journal of Cognitive Neuroscience, 26*, 1736-1747.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime 1.0. Pittsburgh, PA: Psychological Software Tools.

Scott-Van Zeeland, A. A., McNealy, K., Wang, A. T., Sigman, M., Bookheimer, S. Y., & Dapretto, M. (2010). No neural evidence of statistical learning during exposure to artificial languages in children with autism spectrum disorders. *Biological Psychiatry*, *68*(4), 345-351.

Semel, E., Wiig, E. H., & Secord, W. A. (2003). Clinical Evaluation of Language Fundamentals (CELF-4). *The Psychological Corporation*. San Antonio, TX: A Harcourt Assessment. Inc.

Servan-Schreiber, E., & Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(4), 592.

Shi, R., Cutler, A., Werker, J., & Cruickshank, M. (2006). Frequency and form as determinants of functor sensitivity in English-acquiring infants. *The Journal of the Acoustical Society of America*, *119*(6), EL61-EL67.
Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, *81*, 105-120.
Siegelman, N., Boegaerts, L., & Frost R. (2016). Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behavioural Research Methods*, 1-15.
Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., & Frost, R. (2018). Linguistic entrenchment: Prior knowledge impacts statistical learning performance. *Cognition*, *177*, 198-213.
Sound Forge Audio Studio (2013). *Sony Creative Software Inc*.
SpectraPlus software (2008). *Pioneer Hill Software.*
Thiessen, E. D. (2017). What's statistical about learning? Insights from modelling statistical learning as a set of memory processes. *Philosophical Transactions of the Royal Society of London B*, *372*(1711), 20160056.
Thiessen, E. D., & Saffran, J. R. (2007). Learning to learn: Infants' acquisition of stress-based strategies for word segmentation. *Language Learning and Development*, *3*(1), 73-100.
Tomblin, J. B., Mainela-Arnold, E., & Zhang, X. (2007). Procedural learning in adolescents with and without specific language impairment. *Language Learning and Development*, *3*(4), 269-293.
Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, *40*(6), 1245-1260.
Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General, 134*, 552-564.
Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience*, *21*(10), 1934-1945.
Ullman, M. T. (2001). A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience*, *2*(10), 717.
Ullman, M. T., & Gopnik, M. (1999). Inflectional morphology in a family with inherited specific language impairment. *Applied Psycholinguistics*, *20*(1), 51-117.
Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex*, *41*(3), 399-433.
Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of *p*-values. *Psychonomic Bulletin & Review*, *14*(5), 779-804.
Wechsler, D. (2003). *Wechsler intelligence scale for children-WISC-IV*. Psychological Corporation.
Weismer, S. E., Tomblin, J. B., Zhang, X., Buckwalter, P., Chynoweth, J. G., & Jones, M. (2000). Non-word repetition performance in school-age children with and without language impairment. *Journal of Speech, Language, and Hearing Research*, *43*(4), 865-878

Chapter 4: Individual variation in the time course of statistical word segmentation: An ERP investigation

4.1 Introduction

If the ability to detect and learn the distributional regularities within language underlies some aspects of language acquisition (e.g., Saffran, Aslin, & Newport, 1996), then certainly, it would be advantageous for listeners to rapidly detect these patterns. However, existing research has not empirically quantified how quickly learners become sensitive to the regularities embedded within language. A number of computational models of word segmentation, including PARSER (Perruchet & Vinter, 1998), TRACX (French, Addyman, & Mareschal, 2011), TRACX2 (French & Cottrell, 2014), and iMINERVA (Thiessen & Pavlik, 2013) demonstrate that statistical learning involves the extraction of chunks or exemplars from the input and the integration of the regularities across these exemplars over repeated exposure. However, empirical data demonstrating these processes is currently lacking. A further limitation of much of the empirical data on statistical learning is that it relies on behavioural outcome measures following a learning phase, which cannot shed light on the process of learning as it unfolds. Furthermore, these behavioural outcome measures rely on explicit recognition of implicitly learned material, which may underestimate the amount of knowledge accrued in a statistical learning task (e.g., Batterink, Reber, Neville, & Paller 2015). An alternative approach is to use an online measure of statistical learning that can implicitly measure the learner's sensitivity to the distributional regularities within an artificial language. Some neuroimaging work, specifically, the measurement of event-related potentials (ERPs), has been used to implicitly measure the outcome of statistical learning. This work has described event related potentials (ERPs) marking statistical learning, including the N100 (Sanders, Newport, & Neville, 2002), N400 (Cunillera, Toro, Sebastián-Gallés, & Rodríguez-Fornells, 2006; Sanders et al., 2002), P200 (Cunillera et al., 2006; De Diego-Balaguer, Toro, Rodríguez-Fornells, & Bachoud-Lévi, 2007), the P300 (Batterink, Reber, & Paller, 2015; Batterink, Reber, Neville, & Paller, 2015), and the late positivity component (LPC) (Batterink & Neville, 2013). In a novel approach, the present study measured ERPs online during a statistical word segmentation task in order to measure dynamic changes in neural responses over time to provide clear evidence on the processes involved in statistical learning.

Computational models provide a useful framework for understanding the dynamic processes involved during statistical language learning. One of the first examples of this was the PARSER model (Perruchet & Vinter, 1998), which is a memory-based model of word segmentation involving both the extraction and integration of representational units, or chunks. In this model, sensory primitives such as syllables are first experienced within the focus of attention and are then stored in memory as a new representational unit. These representations are reinforced within memory over multiple encounters, while those units that are not re-experienced rapidly decay. These initial percepts or units guide perception, and subsequently-encountered units are entered in a recursive fashion. This assertion naturally fits with the role of statistical learning in language acquisition. For instance, if a group of syllables are extracted as a candidate word, they are more likely to be re-encountered if they are, in fact, a word. On the other hand, if the syllables straddle a word boundary, they are unlikely to be re-encountered and are subject to decay. Additionally, if an element within a chunk occurs within a different chunk, the previously stored chunk experiences interference and loses a degree of activation. Thus, the system rapidly converges towards the words. Taken together, this model clearly stipulates that both the extraction of representational units or words, and the integration of regularities across these units are both necessary processes for statistical learning.

In a related memory-based model of statistical learning, Thiessen and Pavlik's (2013) iMINERVA model demonstrates how memory-based processes can yield sensitivity to structure in a distributional statistical learning task. In this model, stimuli activate prior exemplars already stored in memory as a function of their similarity: The more similar to the exemplar, the stronger the activation for that exemplar. If more than one similar exemplar exists, the stimulus is integrated with the exemplar with the greatest degree of activation. If no similar prior exemplar exists, the stimulus is stored as a new exemplar within memory. Consistent features across multiple stored exemplars become strengthened with repeated activation, while features that are inconsistent are gradually weakened. These combined processes of activation, integration, and decay form the general process by which the model comes to represent a set of exemplars that are prototypical in nature. Unlike PARSER's chunking framework, iMINERVA is not explicitly applied to statistical word segmentation. However, iMINERVA can account for the distributional learning of statistical regularities such as phoneme distributions (e.g., Maye, Werker, & Gerken, 2002), non-adjacent dependencies (e.g., Gómez, 2002), and cue-weighting (e.g., Thiessen & Erickson, 2013). What is important is that both PARSER and iMINERVA stipulate initial chunking or storage of a stimulus or set of exemplars, with subsequent re-activation over repeated presentations preceding eventual storage within memory. Taken together, these models specify how the separable memory-based processes of activation, integration, and decay are involved within a statistical learning paradigm.

While the computational models stipulate how the memory-related processes should be involved in a statistical word segmentation task, the empirical data has yet to clearly demonstrate this. First and foremost, this is due to a limitation in how learning is assessed within a word segmentation paradigm. Overall, studies on word segmentation share an intrinsic limitation: Computations during the learning phase are often only inferred following a lengthy period of language exposure. Typically, participants are asked to assess the familiarity of a trained and foil nonsense word in a two-alternative forced-choice (2AFC) discrimination test only after language exposure (e.g., Saffran, Newport, & Aslin, 1996). There are a number of problems with respect to the 2AFC test, as described by Siegelman, Bogaerts, and Frost (2016). First, testing learning only after exposure runs in stark contrast to the theoretical assumptions of statistical learning, which has been defined as a *process* of continuous assimilation of regularities (e.g., Reber, 1967), with behavioural changes happening incrementally over time. As described by Siegelman, Bogaerts, Christiansen, and Frost (2017) an inherent problem of offline measures is that they are tracing learning retroactively, and may not be doing so in an accurate way. Second, an explicit test is being used to measure an implicit learning process (Siegelman et al., 2016): Offline measures such as a 2AFC require participants to make an overt behavioural response, and to decide between two related stimuli. Given these constraints, it is difficult to disentangle unrelated effects due to encoding and memory capacities, as well as decision-making biases. Finally, the limited number of training items, often only 4 (e.g. Saffran, Aslin, & Newport, 1996) or 6 (e.g., Saffran, Newport, & Aslin, 1996), places inherent limitations on the number of test items. This constraint is made more problematic as accuracy on the 2AFC can only be measured using group-averaged statistics; examinations of individual differences are precluded due to the lack of test sensitivity. Clearly, the use of 2AFC tests obfuscate our understanding of the processes of statistical word segmentation.

In an effort to overcome the limitations of the 2AFC test, Siegelman et al. (2016) used an online measure of statistical learning to measure the rate of learning during a visual sequence learning task. In order to measure online learning, they constructed a self-paced reaction-time (RT) task. For this task, participants were asked to advance the shapes, one at a time, within the sequenced stream at their own pace. The stream consisted of triplet sequences, which were grouped based on the high transitional probabilities between adjacent shapes within a sequence. They found faster RTs for advancing shapes in a predictable compared to a non-predictable stream. Additionally, they found significant RT gains following seven or eight repetitions of the triplets, which clearly demonstrated rapid learning of the sequential structure of the shape stream. Critically, they showed that this RT measure was a valid measure of learning as it correlated with the well-established offline learning measure, the 2AFC. In another examination of changes in behavioural responses over the learning phase, Batterink and Paller (2017) exposed participants to a continuous speech stream composed of nonsense tri-syllabic words, similar to the Saffran, Newport, and Aslin (1996) paradigm, and measured learning online via a target detection task. In the target detection task, participants were asked to monitor for a target syllable, and the target syllable occurred as either the first, second, or third syllable within the nonsense words. By the second presentation of a word, a robust RT effect was elicited such that targets in a predictable position (i.e., a word-final position) were responded to faster than targets in an unpredictable position (i.e., a word-initial position). These RT gains provided further evidence of rapid statistical learning, with faster responding to predictable syllables after only a single word presentation.

Both Siegelman et al. (2016) and Batterink and Paller (2017) found a significant correlation between an offline behavioural test of statistical learning and their online RT measure. However, a significant correlation between online and offline measures has not always been found. In a similar RT paradigm to Batterink and Paller (2017), Batteirnk et al. (2015) measured RTs in a target detection task and ERP responses to the artificial language stimuli as online learning outcomes, and found no correlation between the post-exposure 2AFC task and their online measures. This finding emphasized the implication that explicit knowledge as measured by a 2AFC test may underestimate the implicit knowledge accrued during a statistical language learning task, a limitation previously described by Siegelman et al. (2016). Given this, it is worthwhile considering other implicit measures that more accurately reflect the implicit knowledge acquired during statistical learning.

One possible approach is to examine statistical learning using an implicit measure, such as electroencephalography (EEG) to record ERPs in a statistical learning task. Due to the high temporal sensitivity of ERPs, they are well-suited to examine fine-grained responses to linguistic stimuli. In one of the first examinations of the ERP correlates of word segmentation, Sanders et al. (2002) exposed participants to an artificial language, similar to the Saffran, Newport, and Aslin (1996) paradigm, and measured ERP responses to words before and following language exposure. Sanders et al. (2002) analyzed their ERP data comparing those with high versus low scores on a 2AFC task, and found a larger N100 for word onsets after training only for those with high 2AFC scores, and a greater N400 response to word onsets for both groups. The authors interpreted the elicitation of the N400 as being consistent with the conventional interpretation of the N400 as a marker of lexical and semantic processing (Kutas & Federmeier, 2011), and as supporting some work from the statistical learning field suggesting that segmented words are treated as proto-lexical traces (Saffran, 2001; Graf-Estes, Evans, Alibali, & Saffran, 2007). In a related study, Sanders and Neville (2003) found that word-initial syllables elicited a larger N100 component than word-medial syllables when other sensory cues including loudness, length, and phonemic content were equated. Taken together, these findings helped form the initial interpretation of the N100 as a marker of word segmentation.

Corroborating the findings from Sanders et al. (2002), Cunillera et al. (2006) also found evidence of an N400 component for word-onsets following training in a word segmentation task. Compounded with the previous evidence (e.g., Saffran 2001; Sanders et al., 2002), the authors suggested that the elicitation of an N400 points to the formation of a proto-lexical trace during the word segmentation task. When stress cues marking word onsets were added to the artificial language stream, both the N400 and a P200 were elicited in response to word onsets (Cunillera et al., 2006). Although not pertinent to studies examining word segmentation based on syllable-level distributional regularities, the elicitation of the P200 in response to word-onset cues may have indexed fast auditory learning, and a greater recruitment of neural populations within the auditory cortex (Allison, Wood, & McCarthy, 1986), specifically in response to learned acoustic components within the artificial language. A critical difference between the Cunillera et al. (2006) and the Sanders et al. (2002) study is that the former examined ERP components online during the artificial language learning task. However, the measurement of learning “online” in the Cunillera et al. (2006) study was, in fact, the averaging of ERP responses over 6-minute exposure blocks. Averaging over this duration makes it difficult to identify any dynamic changes in ERP components related to accumulated language exposure. Reconciling this, Cunillera et al. (2009) measured changes in ERP responses to word-onsets over 2-minute exposure blocks. Clear evidence was found for an N400 response to word onsets after only two minutes of artificial language exposure. Supporting the electrophysiological findings, participants could behaviourally identify newly segmented words after one minute of language exposure. What was interesting was that the N400 decreased in amplitude with longer exposure. This result shed new light on measuring dynamic changes within statistical learning, and further strengthened the interpretation of the N400 as a marker of proto-lexical trace formation in statistical learning tasks.

Measuring the dynamic changes during a statistical segmentation task gained attention following the work from Cunillera et al. (2006) and Cunillera et al. (2009). Abla, Katahira, and Okanoya (2008) had participants listen to a non-linguistic auditory tone sequence, the elements of which were three-tone “words”, and was consistent with the regularities within the artificial language stimuli used by Saffran, Newport, and Aslin (1996). Participants were exposed to the auditory input for three 6.6-minute sessions. Following exposure, participants performed a behavioural test in which they identified familiar tone sequences from the training stimuli. Based on this behavioural test, participants were divided into three groups: High, middle, and low learners. A comparison of ERPs between the learner groups revealed that for the high learners, three-tone “word” onsets elicited N100 and N400 components within the first 6.6-minute listening session, with these components being elicited for the middle learners in the second session, and with no reliable N100 or N400 effect for the low learners. Similar to related studies (Cunillera et al., 2006; Cunillera et al., 2009; Sanders et al., 2002), the authors interpreted the N400 effect as reflecting on-line segmentation.

Findings of a difference in ERPs between good and poor statistical learners (Abla et al., 2008; Sanders et al., 2002) raises the possibility that individual variation in statistical learning ability may influence online segmentation. The examination of statistical learning of auditory sequences in musicians offers an interesting examination of how domain-specific expertise influences the processing of structured input. Musicians are a unique population in which to examine statistical segmentation, as previous research has shown they are better at grouping and processing complex auditory patterns (Boh, Herholz, Lappe, & Pantev, 2011; van Zuijen, Sussman, Winkler, Näätänen, & Tervaniemi, 2004). Additionally, research has shown that musicians demonstrate faster working memory updating than non-musicians (George & Coch, 2011), which may be a necessary process for the extraction and consolidation of statistical patterns, as shown in Chapter 2 of this thesis, and by others (e.g., Cunillera et al., 2009; Lopez-Barroso, de Deigo-Balaguer, Cunillera, Camara, Münte, & Rodríguez-Fornells, 2011; Palmer & Mattys, 2016). Closely related to word segmentation paradigms, the findings from tone “word” segmentation paradigms in musicians have shown N100 and N400 triplet onset effects (Vasuki, Pragati, Sharma, Ibrahim, & Arciuli, 2017). Additionally, ERP markers were found for musicians segmenting a sung language, including a reduction in P200 amplitude over increased exposure (François & Schön, 2011), and modulation of the N400 for word-onsets (François, Jaillet, Takerkart, & Schön, 2014), with expert learners showing an inverted U-shaped curve in N400 amplitude over the language exposure period. These findings generally point to the conclusion that expertise in a particular domain seems to influence neural responding during segmentation. Furthermore, this individual variation may point to some of the cognitive processes involved in segmentation. Specifically, the inverted U-shaped amplitude modulation demonstrated by François et al. (2014) was attributed to the consolidation of the extracted phonological units into templates or exemplars. Moreover, the expert effects exhibited across these studies are consistent with other "high learner" or expert learner effects in the statistical language learning literature (e.g., Abla et al., 2008; Sanders et al., 2002).

From this body of evidence, it seemed that the N100 and N400 were reliable markers of statistical learning, reflecting word segmentation and pre-lexicalization processes, respectively. However, these effects have not consistently been reported. Cunillera et al. (2006) did not observe an N100 effect, but rather reported a P200 increase for stressed syllables embedded in a structured artificial language, but not an unstructured stream. Additionally, the N100 and the P200 were not observed in response for an unstressed structured stream. However, it is important to note that the inclusion of stress markers indicating word onsets adds an additional linguistic cue beyond the transitional probabilities of syllables, which makes it difficult to interpret these results in terms of the ERP components related to the extraction or integration of purely conditional relationships between syllables.

In a continuation of the investigating of ERP markers for word segmentation, De Diego-Balaguer and colleagues (2007) found that listeners elicited a P200 in response to words in an artificial language by the 3rd minute of exposure. The modulation of the P200 was interpreted as a "reorganization" of the structural information. That is, the P200 was considered indicative of chunking of the speech signal prior to the extraction of its underlying structure. This suggestion is consistent with interpretations of the P200 as reflecting perceptual segregation of auditory input (Snyder, Alain, & Picton, 2006). De

Diego-Balaguer and colleagues (2007) also found evidence of an N400, which was related to word learning following language exposure. This pattern of findings is interesting, as it seems to disentangle both the extraction of structural rules as indicated by the modulations in the P200, and successful word learning as marked by the N400. Taken together, the findings from Cunillera et al. (2006) and de Diego-Balaguer (2007) may indicate ERP that reflect perceptual segregation or chunking when attending to structured auditory input, specifically the P200.

The lack of N100 and N400 effects has also been reported by Batterink, Reber, Neville, and Paller (2015) and Battink, Reber and Paller (2015). In both studies, ERP responses to trained nonsense words were recorded following exposure to an artificial language (e.g., Saffran et al., 1996). Trained words elicited an enhanced late positive potential (LPC) following language exposure, which was thought to be indicative of explicit knowledge for the newly segmented words. The LPC is generally found over parietal scalp sites, begins around 400-500ms after the onset of a stimulus (Paller & Kutas, 1992; Rugg & Curran, 2007), and is often associated with explicit memory for repeated stimuli (Rugg et al., 1998). Batterink et al. (2015) and Batterink, Reber, and Paller (2015) also examined the amplitude of the P300 component during a target detection task following language exposure. For this task, participants were asked to monitor for a specific syllable within a stream of nonsense words from the artificial language. The amplitude of the P300 was found to differ as a function of syllable position, with word-initial syllables eliciting the largest P300, word-medial syllables eliciting a moderate P300, and word-final syllables eliciting the smallest P300. This finding is consistent with interpretations of the P300 as associated with attention and memory operations (Polich, 2007), and reflecting allocation of attentional resources to a target stimulus (Polich, 2003; 2007). Specifically, targets that are less probable, such as word-initial syllables, are thought to elicit a larger P300 response. Conversely, targets that are more probable, such as word-final syllables, should elicit a smaller P300 response. This decrease in P300 amplitude as a function of syllable predictability, based on the statistics within the nonsense language, may reflect a facilitation in processing for predictable syllables. Relatedly, Batterink et al. (2015) found that those who were explicitly trained on the nonsense words prior to language exposure showed a larger and earlier P300 to word-final syllables than those with no training. They speculated that this larger anticipatory P300 for the explicitly trained participants reflected anticipatory processing and greater overall processing effort for predictable syllables.

4.1.1 Present study

Although much of the work examining the ERP components related to statistical language learning is convincing of extant ERP markers for statistical word segmentation, a consistent shortcoming is that ERP responses are averaged over long periods exposure time, with analyses being focused on blocks of 6 or more minutes (e.g., Abla et al, 2008; Cunillera et al., 2006; Sanders et al., 2002), or examined only following language exposure (e.g., Batterink, Reber, & Paller, 2015; Batterink, et al., 2015). Because of this, it is difficult to clearly distinguish the extraction and integration processes that are thought to underlie statistical word segmentation (e.g., Perruchet & Vinter, 1998; Thiessen & Pavlik, 2013). To overcome this, the present study examined the dynamic changes in ERP components online during a statistical word segmentation task.

In the present study, participants were exposed to a structured, artificial language for 21-minutes, with ERPs recorded for the duration of the language exposure period. Following language exposure, participants completed the conventional 2AFC word identification task. It was of interest to explore individual variation in the ERP responses based on statistical learning abilities, so the sample was divided into high- and low-learners based on the 2AFC scores, and ERPs during the exposure period were compared between the learner groups. Following the results of de Diego-Balaguer (2007), Cunillera et al. (2006), and François and Schön (2011), analyses were focused on modulations in the P200 component over exposure time, as this component may be indicative of extraction of the statistical regularities within the structured artificial language. Also, given the syllable-position effects from Batterink et al. (2015) and Batterink, Reber, and Paller (2015), analyses of the P200 were focused on responses to maximally predictable syllables, that is, word-final syllables.

Two main hypotheses follow from the analyses of the continuous ERP data. First, it was hypothesized that if participants were actively segmenting the artificial language based on the transitional probabilities embedded within the speech stream, responses to predictable syllables should be modulated as a function of accumulated exposure. Second, if the modulation effect of the P200 differed across learner groups, this would provide evidence of a relationship between an implicit measure of extraction of the statistical regularities and the outcome of statistical language learning. Although an assessment of the evolution of ERP responses to predictable syllables was the primary interest of the present study, it was also worthwhile considering the N100 and N400 effects marking word segmentation initially reported by Sanders et al. (2002). To do this, pre- and post-training effects were approximated by examining N100 and N400 amplitude in the first (minute 1) and final (minute 21) minutes of the language exposure periods, both at the group level, and comparing these effects between the high and low learner groups.

4.2 Method

4.2.1 Participants

Twenty-two participants (14 female) were recruited from The University of Western Ontario community for this experiment. All participants were between 18 and 30 years old (*M* = 20.24, *SD* = 2.70), native English-monolingual speakers, and reported no hearing impairments. Participants either received course credit or were paid $10/hour for study participation. Ethics approval for all study procedures and materials was obtained from the University of Western Ontario Non-Medical Research Ethics Board, and written informed consent was obtained from all study participants.

4.2.2 Stimuli

4.2.2.1 Artificial language stimuli

The artificial language was based on the stimuli described by Saffran, Newport and Aslin (1996). The language was composed of an inventory of 12 CV syllables, combined to create six trisyllabic "words": *patubi, tutibu, babupu, bupada, dutaba, pidadi*. Transitional probabilities of syllables ranged from 0.33 to 1.0 within-word, and from 0.1 to 0.2 across word boundaries assuming equal distribution of each word preceding and following all others.

The artificial language was constructed from audio recordings of a female native-English speaker using a neutral vocal effort. Recordings of the speech stimuli were made in a double walled IAC sound booth with a pedestal microphone (AKG C 4000B) located approximately 30cm from the speaker's mouth and routed to a USBPre 2 pre-amplifier (Sound Devices) using SpectraPlus software (Pioneer Hill Software, 2008). Recordings were made of each of the 12 target syllables in the middle of a three-syllable sequence, within every co-articulation context required for the language. Eight repetitions of each sequence were recorded, and the token with the most neutral pitch contour and best sound quality was chosen and uploaded into Sound Forge Audio Studio (Sony Creative Software, *version* 10.0) editing software. Middle syllables from the recorded tokens were extracted by identifying the final offset of vowel oscillation in the previous syllable to the offset of vowel oscillation in the target syllable. These were then concatenated to create the final 21-minute stream of words. The stream consisted of 360 tokens of each word in random order, with no word presented twice in sequence. The language maintained a consistent speech rate (average 5.1 syllables/s) using a time stretch, and was normalized to a pitch of F0 = 196 Hz using the pitch shift in Sound Forge Audio Studio. There were no pauses between words; as such, the only cues to word boundaries were the lower transitional probabilities for between-word syllable pairs.

4.2.2.2 Test phase stimuli

An additional six non-word foils were constructed from the same 12 CV syllables as the artificial language, but which were not included in the training set: *pubati, tapudi, dupitu, tipabu, bidata, batipi*. Non-word foils were created with the constraint that within-word transitional probabilities would be zero. Syllables were drawn from the same recording inventory as the artificial language stimuli, with appropriate co-articulation contexts. Note that we used fully new non-word foils with transitional probabilities of 0 based on previous language exposure, rather than trisyllabic part-word foils consisting of two syllables from a trained item plus an incorrect syllable (e.g., Saffran, Newport, & Aslin, 1996). Discrimination accuracy at test is generally higher when fully new non-word foils are used rather than part-words. As our primary goal was not to discover if statistical learning of language was present generally, but rather, whether we could uncover neural indices of segmentation during and following language exposure, the more sensitive measure was used.

4.2.3 Procedure

4.2.3.1 Artificial language exposure phase.

The artificial language was played over speakers at a comfortable listening volume. Participants were told they would hear a nonsense language, with no information provided about the length or the number of words within the language. These deliberately vague instructions minimized the chance of participants trying to explicitly learn the language's structure during the experiment. To reduce eye movements, participants were instructed to look at a fixation cross during the exposure phase while during EEG recording. Every seven minutes during the exposure phase, they received a short break. No other task was presented to the participants during the exposure phase.

4.2.3.2 Test phase

Following the exposure phase, participants completed a two-alternative forced-choice (2AFC) task delivered by E-Prime 2.08 (Schneider et al., 2002). For each test item, participants heard a trained word from the artificial language paired with a non-word foil, separated by 500ms of silence. Presentation order of trained words and non-word foils were randomized across trials. Each non-word foil was paired exhaustively with each trained word, comprising 36 total test pairs, and presented in a fixed random order. Subjects were instructed to indicate which word "sounds more like something you heard in the language", and to select *"A"* or "*L*" on the keyboard to indicate the first or second stimulus, respectively. Instructions remained on the screen throughout the test phase. Behavioural accuracy on the task was calculated for each participant as the percent of correct identifications of trained words.

4.2.4 EEG recording and pre-processing

Continuous EEG was acquired during the exposure and test phases using a BioSemi ActiveTwo system consisting of amplifier-embedded Ag/AgCl electrodes at 32 scalp sites (Fp1/2, AF3/4, F7/8, F3/4, T7/8, C3/4, CP5/6, CP1/2, P7/8, P3/4, PO3/4, O1/2, Fz, Cz, Pz, and Oz) and two mastoid sites per the International 20-30 system. Electrooculogram (EOG) was recorded from four active electrodes placed on the outer canthus of either eye, and above and below the left eye. A Common Mode Sense active electrode and a Driven Right Leg passive electrode were used as ground. Data were acquired at a 512 Hz sampling rate, filtered online at 0.1-100 Hz bandpass and 60Hz notch filters, and impedances kept below 20 kΩ.

Due to the use of naturally produced speech in our artificial language stream, all syllables were of varying length. We were interested in recording ERPs at the onset of each syllable during the exposure phase, however, stimulus presentation software can introduce inaccuracies in EEG time-locking due to latencies introduced by buffering the acoustic stimulus through a digital-to-analog converter prior to outputting the signal to a speaker. To eliminate this problem, time-locking of the ERPs to the auditory stimuli was achieved by directly tracking the onset of each syllable in the stream and encoding it alongside the continuous EEG data using an auxiliary analog-to-digital interface (StimTracker; Cedrus Corporation, 2010). To do this, the left audio channel of the digital audio stream contained a 10ms click, manually aligned to the onset of each auditory syllable within the right audio channel. The left and right audio outputs were electrically isolated so that the participant heard only the right channel, played through a speaker, whereas the left channel was fed to the StimTracker. The EEG triggers produced by the Stim Tracker were used to mark syllable onsets throughout our analyses.

ERP data were processed in MATLAB (2015b) using the EEGLAB (*version* 13; Delorme & Makeig, 2004) and ERPLAB (*version* 5.1.1.10; Lopez-Calderon & Luck, 2014) toolboxes. Continuous EEG was bandpass-filtered from 0.1 to 30 Hz, and referenced to the algebraic average of the left and right mastoid electrodes. Scalp EEGs were then submitted to Independent Component Analysis (ICA) using the extended *fastica* routine in EEGLAB (Hyvärinen & Oja, 2000) to identify and remove ocular and exogenous channel-wise artifacts. Epochs were time-locked to the onset of each syllable within the language exposure and test phase, (-50 to 800ms post stimulus onset), and baseline corrected to the 50ms pre-stimulus interval. Following epoching, data were submitted to moving-window (200ms) peak-to-peak artifact detection to identify and remove any remaining artifacts greater than 100uV. Three sets of epochs were created reflecting the onset of each syllable within a word (onset, medial, and final); these were further grouped into separate 1-minute average bins reflecting the mean syllable-wise ERP response during each minute of exposure.

4.2.5 Data analysis

First, performance on the 2AFC behavioural measure was analyzed. Participants scores were calculated as the number of correct identifications of a trained word within test pairs. For our subsequent analyses, participants were divided into two groups based on a median split of 2AFC accuracy scores. This created “low-learner” and “high-learner” groups, which was our grouping variable for the analysis of ERP responses during the exposure phase.

Next, the P200 component in response to word-final syllables throughout the exposure phase was examined. On the basis of visual inspection, the peak of the P200 was identified at 250ms post-stimulus onset, and average amplitude of the amplitude was computed from 220-280ms (i.e., 250ms +/- 30ms) for all syllables within the word-final position. Next, the mean P200 voltages were averaged together across each one-minute time interval to create an average voltage for the P200 for word-final syllables for each minute during the language exposure phase.

The mean amplitude of the word-final P200 response was analyzed using a Growth Curve Analysis (GCA) technique to compare the amplitude of the component between the high- and low-learner groups. GCA is a multilevel regression technique used to model and compare time-varying data. A significant advantage for GCA over univariate techniques is that it does not require multiple comparisons at discrete time points or averaging across large time bins, both of which can obscure important temporal effects. This technique can also incorporate both group-level and individual-level differences into the model. Here, the time variable was modeled using an orthogonal polynomial, which captures the curvature of the data and is orthogonally transformed to avoid collinearity of higher order polynomials. Groups were compared by adding group-level differences, stepwise, to each polynomial term of the GCA model, and using analysis of variance (ANOVA) to compare improvements in model fit. Additionally, the parameter estimates are compared between groups to examine differences in the shape of the data. Significant differences in the intercept term represent different means, while differences in the linear term of the model are associated with different slopes. The quadratic and cubic terms of the model represent its curvature, and significant differences between these terms can be interpreted as different quadratic and cubic curvatures between groups. As higher-order polynomials are non-asymptotic (their values do not have plateau-like sections), they do not provide an adequate fit for asymptotic data. To resolve this difficulty, we restricted the length of the tail data that were included in our analyses (Mirman, 2014). Using visual inspection, it was clear that the data reached asymptote across both groups by the 4th minute of exposure. Thus, analyses were constrained to ERP responses from minute 1 to minute 4 during the artificial language exposure phase [4]. All GCA analyses were carried out in *R* (*version* 3.3.1; R Core Team, 2016), using the *lme4* and *lmerTest* packages.

Electrodes were grouped into regions of interest (ROIs) arranged in a 3x3 grid over the scalp (left/midline/right; anterior/central/posterior), and average voltages from the electrodes within each ROI were averaged together to create a mean amplitude within each ROI for the selected time-window entered into the analyses. For analysis of the ERP responses during the exposure phase, we examined responses for all frontal electrodes across left, midline, and right hemispheres.

Given that an N100 and N400 have been cited as markers of word segmentation (Abla et al, 2008; Abla et al., 2009; Cunillera et al., 2006; Sanders et al., 2002), the presence of these components was investigated in the present dataset. Because the artificial language used in the present study most closely resembled that used in Sanders et al., (2002), both in the vocabulary of the language and the exposure duration, their analyses were replicated here. In their study, Sanders et al. measured the N100 and N400 component pre- and post-training. Pre- and post-training measures were not collected in the present study, as responses to the artificial language were measured continuously throughout language exposure, so ERP responses to word onsets averaged across the 1st minute and 21st minute were analyzed. The logic behind selecting these time points is that

[4] To avoid introducing experimenter bias into the selection of the 4-minute time window, parallel analyses were conducted using different time windows (e.g., 3 minutes, 5 minutes, 7 minutes), and verified that the results were the same across these similar models. In doing so, selection of the 4-minute tail was approached in an unbiased manner.

it should approximate any effects related to pre- versus post-training that were uncovered in the Sanders et al. paradigm. Furthermore, examining ERP responses at the 21st minute of exposure is roughly equivalent to the total accumulated language exposure for participants in the Sanders et al. study. Consistent with Sanders et al., the average amplitude of the N100 was measured between 70 and 130ms post stimulus-onset, and the average amplitude of the N400 was measured between 200 and 500ms post stimulus-onset. Modulation of the N100 and N400 due to training were analyzed both at the group level, and between high- and low-learner groups. For these analyses, the criteria for selecting the high- and low-learner groups was the same as the grouping criteria for the GCA analyses.

As with the GCA analyses, electrodes were grouped into regions of interest (ROIs) arranged in a 3x3 grid over the scalp (left/midline/right; anterior/central/posterior), and average voltages from the electrodes within each ROI were averaged together to create a mean amplitude for each ROI for the selected time-window were entered into the analyses. For the supplemental analyses, ROIs were treated as repeated measures to examine the topographical distribution of the N100 and N400 effects. Pairwise post-hoc *t*-tests were used to identify ROIs for which the condition-wise effect was significant.

4.3 Results

4.3.1 2AFC behavioural outcome measure

Scores on the 2AFC measure (M = 69.79%, SD = 9.7%, min = 52.77%, max = 80.56%) were significantly greater than chance, $t\ (1, 23) = 9.9929$, $p < .001$, where chance was defined as identifying trained words correctly on 50% of test trials. A scatterplot of participants' performance on the 2AFC measure is presented in Figure 4.1.

For the growth-curve analysis of ERPs during the exposure phase, a median split of 2AFC scores was used to divide the sample into "low-learners" and "high-learners". Three participants scored at the median value (69.44% correct), and were grouped as low learners in order to keep the group sizes roughly equivalent. Thus, the sample was divided into 12 low learners (2AFC scores: $M = 63.42\%$; $SD = 5.86\%$) and 10 high learners (2AFC scores: $M = 79.44\%$; $SD = 2.68\%$).

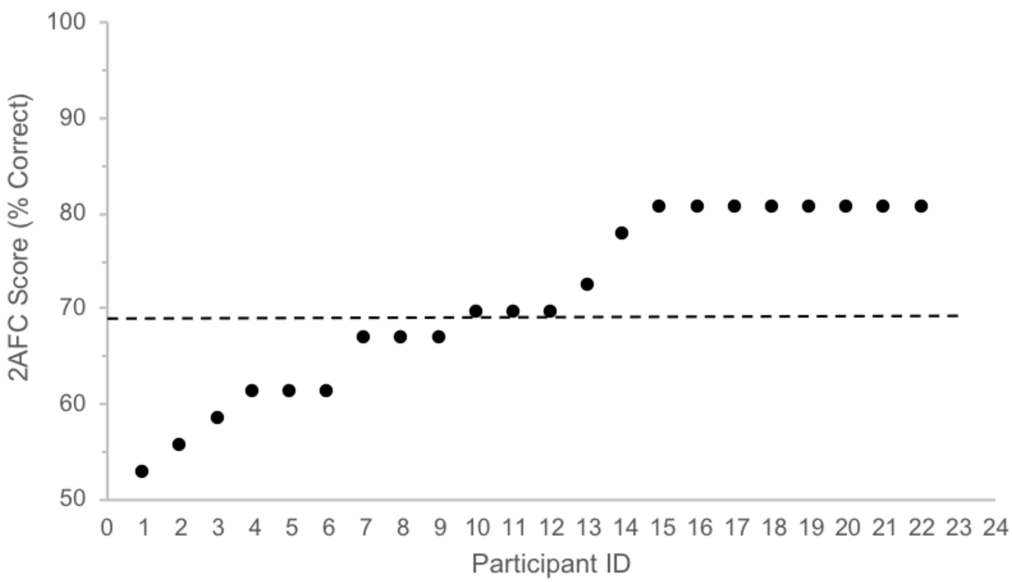


Figure 4.1 Individual performance on the 2AFC discrimination task following language exposure. Scores are reported as percent correct (out of 36 test pairs). Horizontal trendline represents group average score of 69.79%.

4.3.2 Growth curve analysis of exposure phase event-related potentials

Figure 4.2 depicts the amplitude of the P200 in response to word-final syllables over the first four minutes of language exposure for the high- and low- learner groups. The mean amplitude of the word-final P200 response was analyzed using a Growth Curve Analysis (GCA) technique to compare the amplitude of the component between the high-and low-learner groups across language exposure; specifically, the first four minutes of language exposure. Average amplitude for both groups at each time point were entered into the analyses to examine changes in P200 amplitude over time. The overall amplitude curves were modelled with third-order orthogonal polynomials on the time term (minute of exposure) and fixed effects of learner (low/high) on all time terms. The high-learner group was treated as the baseline in the model, and parameters were estimated for the low-learner condition. The model also included random effects of participants on all three orthogonalized time terms. The fixed effects of learner-group (high versus low) were added individually and their effects on model fit were evaluated using model comparisons. Improvements in model fit were evaluated using a -2 times the change in log-likelihood, which has a chi-square distribution with degrees of freedom equal to the number of parameters added.

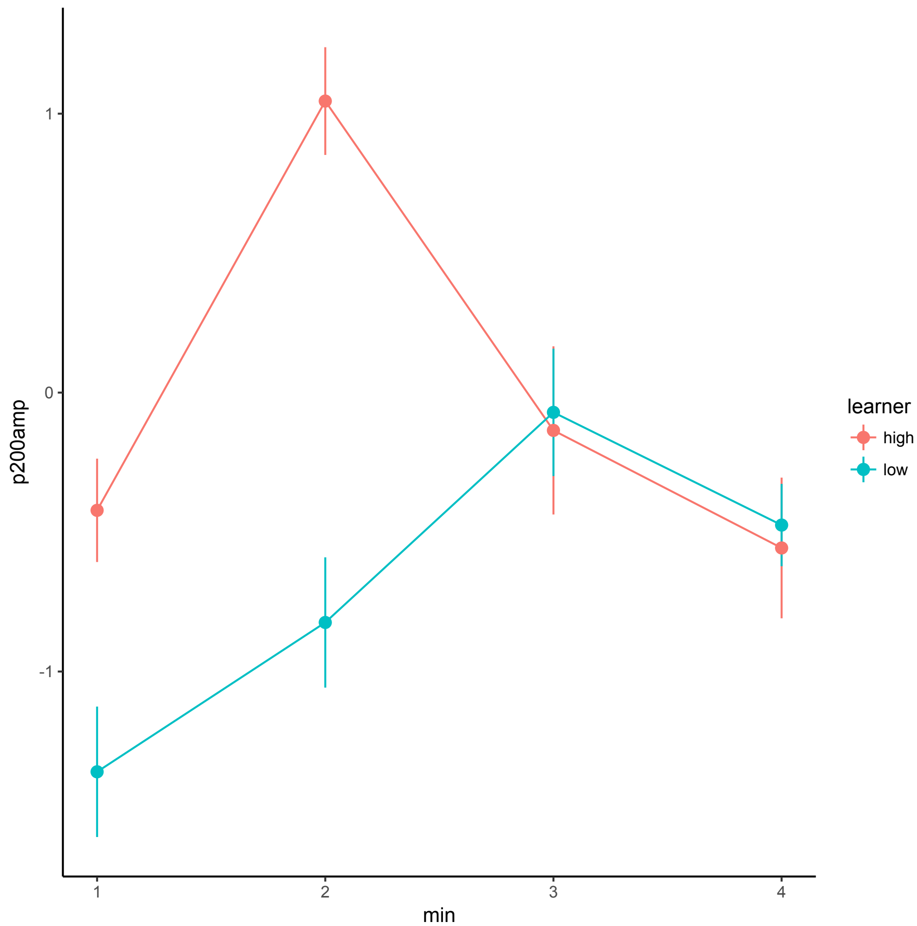


Figure 4.2 Average P200 amplitude across frontal electrode sites for learner groups during artificial language exposure.

The effect of learner group on the intercept did not improve model fit, $\chi^2(1) = 3.27$, $p = .070$, nor on the linear term, $\chi^2(1) = 0.80$, $p = .368$, or the quadratic term, $\chi^2(1) = 3.27$, $p = .070$. The effect of learner group on the cubic term, however, did improve model fit, $\chi^2(1) = 7.51$, $p = .006$, indicating that the low and high learner groups differed in the amplitude of their P200 responses to word-final syllables over the first four minutes of artificial language exposure. Table 4.1 shows the fixed effects parameter estimates and their standard errors, along with *p*-values estimated using the normal approximation for the *t*-values. The significant effect of learner group on time terms indicated that the curvature of P200 amplitude over the first four minutes of the exposure phase were different between groups. There was a significant effect of learner group on the linear term, indicating a steeper, positive linear slope in P200 amplitude for the low compared to the high learners. There was also a significant effect of learner group on the cubic term, indicating a steeper (i.e., sharper) peak in P200 amplitude for the high versus the low learner group.

Table 4.1 Parameter estimates for growth-curve analysis of effect of learner group (high/low) on P200 responses to word-final syllables during artificial language exposure.

	Estimate	SE	*t*	*p*
Intercept	-0.01736	0.25292	-0.069	0.94590
Linear	-0.35452	0.30697	-1.155	0.26053
Quadratic	-0.94438	0.41922	-2.253	**0.03458**
Cubic	0.76189	0.26399	2.886	**0.00858**
Low Learner	-0.66496	0.34245	-1.942	0.06508
Low Learner: Linear	1.11638	0.41564	2.686	**0.01350**
Low Learner: Quadratic	0.47463	0.56762	0.836	0.41204
Low Learner: Cubic	-1.06946	0.35745	-2.992	**0.00672**

Note. Bolded values are significant at the $p < .05$ level.

4.3.3 N100 training effect

A 2 (minute of exposure: 1st, 21st) by 9 (ROI) by 2 (leaner: high, low) analysis of variance was conducted to determine if there was an effect of training on N100 responses to word onsets. A Greenhouse-Geisser correction for multiple comparisons was applied to within-subjects' tests. There was a significant interaction between minute of exposure and ROI, F (8, 160) = 1.772, p = .002. Post-hoc analyses with a Bonferroni correction for multiple comparisons revealed that the amplitude of the N100 was greater in the 1st compared to the 21st minute at left frontal (p = .009), central frontal (p = .031), and right frontal (p = .031) ROIs. The main effect of minute of exposure was not significant, F (1, 20) = 1.310, p = .266, nor were the interactions between minute of exposure and learner, F (1, 20) = 1.612 p = .219, or the three-way interaction between minute of exposure, ROI, and learner, F (8, 160) = 1.772, p = .180. Thus, N100 amplitude was significantly greater at the onset of training, and there was no effect of learner on changes in N100 amplitude over language exposure.

4.3.4 N400 training effect

A 2 (minute of exposure: 1st, 21st) by 9 (ROI) by 2 (learner: high, low) analysis of variance was conducted to determine if there was an effect of training on N400 responses to word onsets. A Greenhouse-Geisser correction for multiple comparisons was applied to within-subjects' tests. There was a significant interaction between minute of exposure and ROI, $F(8, 160) = 2.629$, $p = .049$, however, examination of the simple main effects of minute of exposure at each ROI again revealed no significant differences for both Bonferroni corrected and uncorrected comparisons, $p > .181$, *all cases*. There was no significant effect of minute of exposure, $F(1, 20) = 0.235$, $p = .633$, no interaction between minute of exposure and learner status, $F(1, 20) = 0.494$, $p = .490$, and no three-way interaction between minute of exposure, ROI, and learner status, $F(8, 160) = 1.204$, $p = .395$. Thus, there was no learning effect and no effect of high versus low learners on modulation of the N400 due to language exposure.

4.4 Discussion

In the present study, participants' ERPs were recorded online during exposure to a structured, unsegmented artificial language (e.g., Saffran, Newport, & Aslin, 1996). Participants were grouped based on a post-exposure behavioural measure into "high" and "low" learners, and ERPs during language exposure were compared between these two groups. Analysis of ERPs during language exposure revealed that the high-learner group showed an increase in P200 amplitude in response to word-final syllables at the second minute of language exposure, with this response attenuating with subsequent exposure. The ERPs for the low-learner group, on the other hand, did not show sensitivity to the predictable word-final syllables with accumulated language exposure. The present data also did not show evidence of N100 or N400 responses to trained words from the artificial language, despite these effects being described as markers of word segmentation (e.g., Sanders et al., 2002). These results present evidence of the dynamic nature of word segmentation, and offer a novel approach to measure statistical learning in an online fashion.

Statistical learning can be viewed as a continuous, rapid, and incremental learning process, whereby computations of local statistics in novel environments enable people to predict and more efficiently process incoming input. Evidence for this continuous and rapid learning process may be provided by the dynamic nature of the P200 component in the present study. This finding allows us some insight in determining the underlying cognitive processes engaged in statistical word segmentation. To interpret these findings, it is important to consider not only the conventional interpretation of the P200 within cognitive neuroscience, but also the functional significance of this component in relation to processes believed to be fundamental to statistical learning.

One possible interpretation of the P200 is that it is related to perceptual segregation and attention (Hillyard, Hink, Schwent, & Picton, 1973; Reinke, He, Wang, & Alain, 2003). For instance, Snyder et al. (2006) demonstrated that auditory segregation is dependent on attention. In their study, the amplitude of the P200 auditory evoked response was positively correlated with the perceptual segregation of a continuous and structured stream of tones which followed an ABA pattern. Furthermore, the amplitude of the P200 diminished when participants ignored or did not attend to the ABA pattern. However, it is not just segregation that is essential to word segmentation: Learners must also encode the segmented units or chunks within memory (e.g., Perruchet & Vinter, 1996; Thiessen & Pavlik, 2015). And, attentional resource availability is an important aspect of long-term retention, particularly when specific encoding operations are required (Craik & Byrd, 1982; Craik & Lockhart, 1972; Triesman, 1964). Supporting the interpretation of the P200 as a marker of perceptual learning and encoding, de Diego-Balaguer et al. (2007) found a correlation between the participants' perception of syllable groupings, measured behaviourally, and the magnitude of the P200. The P200 increase was interpreted as a perceptual change due to the re-allocation of attention to structural regularities within the artificial language. Based on these findings relating segmentation and learning, it would be consistent to interpret the elicitation of the P200 component in the present paradigm as a marker of attentional allocation to structured material.

The notion that language learners are encoding the regularities within the artificial language is particularly apparent considering the present study measured responses to predictable, word-final syllables. If learners are, in fact, encoding the regularities of the structured speech stream in memory, it would follow that given the first two syllables of a word (e.g., *tutu-* or *babu-*), the final syllable would be perfectly predicted by the learner based on the transitional probabilities within the language (e.g., *tuti* → *bu* or *babu* → *pu*). Measuring participants' ability to identify words from the language was achieved in the present study by using scores on the statistical learning outcome measure, the two-alternative forced-choice (2AFC) to group participants into "high" and "low" learners, relative to the group-averaged scores. Thus, the emergence of the P200 in response to word-final syllables in the high-learner group may reflect an increase in attentional allocation to syllables consistent with learners' successful segmentation of the speech stream. Such a conclusion is consistent with the findings from Batterink et al. (2015), who measured implicit behavioural and neural responses to word-final syllables, and found that those with better performance on the 2AFC word identification task also showed greater sensitivity in their behavioural and neural measures to word-final syllables. Thus, the online measurement in the present study may highlight the process of segmentation of statistical learning. That is, the extent to which individuals spontaneously detect and learn the distributional regularities within language.

A novel aspect of the present findings is that the amplitude of the P200 component was maximal at the second minute of artificial language exposure for the high-learner group. Similar findings were reported by Cunillera et al. (2009), who found that the amplitude of the N400 response to word onsets was maximal in the second 2-minute exposure block, and the findings from de Diego-Balaguer et al. (2007), who found that the P200 response to word onsets was maximal at the third minute of language exposure. One possible interpretation of the amplitude shift with accumulated exposure is that the listeners had a sufficient amount of information by the second minute in order to segment the language with two minutes of exposure, as they would have heard roughly 34 presentations of each word token in this time (for 6 words in total, at roughly 103 words-per-minute, for two minutes). Indeed, in the original demonstration of statistical learning, Saffran, Aslin, and Newport (1996) found that infants could identify words from an artificial language after two minutes of language exposure. Thus, two minutes of language exposure may be sufficient to extract proto-lexical phonological traces from the structured, unsegmented speech stream. Such an interpretation is consistent with previous work, suggesting that one of the primary outcomes of word segmentation within this paradigm is the formation of candidate words (e.g., Graf Estes et al., 2007; Saffran, 2003; Sanders et al., 2002).

It is important to note that the speech stream in statistical word segmentation studies is, in its essence, devoid of many of the cues extant in natural language, including stress, pause, or prosodic cues, not to mention extra-linguistic context. For this reason, it could be that the P200 component for the high-learner group reflected successful segmentation based on the transitional probability cues, but that additional cues in the speech stream would have facilitated further processing. Moreover, the absence of a neural marker for segmentation in the low-learner group may be reflective of their need for additional cues to successfully segment speech. The notion that the high-learner group may shift their segmentation strategy, or may shift their attention to additional segmentation cues, is consistent with behavioural work showing that listeners shift their strategy from tracking words to uncovering the underlying structure when the signal contains cues that may facilitate this process (Peña, Bonatti, Nespor, & Mehler, 2011). Further, developmental research has shown that when both transitional probability cues and stress cues are present in a speech stream, 7-month-old infants show a preference for the transitional probability cues, while 9-month-old infants show a preference for the stress cues (Thiessen & Saffran, 2003). Based on these findings of a developmental change suggest a shift in the way the speech signal is processed. It may be that examining speech segmentation using combinatorial or iterative cues may help explain the difference in sensitivity to the regularities in the artificial language between the low- and high-learner groups. However, another possible interpretation of the lack of an ERP response in the low-learner group is that they were not appropriately allocating attention to the structural regularities within the artificial language. As discussed above, the P200 in response to word-final syllables for the high-learner group is perhaps reflective of attention to the structured material, and behavioural work has shown that attention is necessary to for word segmentation (Toro et al., 2005). It could be that additional segmentation cues in the speech stream may serve to facilitate attention for the low-learner group.

One surprising finding within the present data was the lack of an N100 and N400 effect for word segmentation, as these have been previously described as "markers of word segmentation" (Abla et al., 20080; Cunillera et al.,2006; Cunillera et al., 2009; Sanders et al., 2002). In fact, there was greater N100 negativity at the pre-training compared to post-training. However, one explanation for the absence of these effects, and the N100 effect at training onset, can be taken from the classical interpretation of these ERP components. First, the auditory N100 is conventionally seen to be modulated by a function of expectation, with greater amplitude for unexpected stimuli (Spreng, 1980), and is generated within primary auditory cortex (Godey, Schwartz, de Graaf, Chauvel, & Liegeois-Chauvel, 2001). It seems unlikely, then, that recognition of phonological information, which is mediated by higher level cortical areas, would be reflective in an N100 modulation. An additional consideration is that there is an insufficient amount of acoustic information presented within 100ms for the listener to recognize a novel word as familiar or unfamiliar, as the first 200ms post stimulus onset is largely dominated by perceptual analysis. Thus, it is unlikely that the N100 is reflective of word segmentation as this interpretation is not acoustically plausible. Second, the N400 is well-described as a response to words or meaningful stimuli (*see* Kutas & Federmeier, 2011), and is reflective of semantic access within long term memory. Thus, it seems unlikely that the recognition of a phonological from an artificial language would generate such a response. Taken together, given how the N100 and N400 are often interpreted in the literature, and the specific acoustic and processing constraints of the N100, it is not surprising that neither component were elicited in the present paradigm in response to trained words.

There are some limitations that need to be considered when interpreting these results. The first is the inherent limitation of measuring ERPs in response to rapidly presented stimuli. The syllables in the language for the present study had a duration of roughly 195ms (or, 5.1 syllables per second), and the computation of the ERPs in response to these rapidly presented stimuli is complicated by baseline issues. Specifically, since was is no silence between successive syllables the baseline voltage correction for ERP responses to each syllable was clearly affected by the ERP to the preceding syllable. To correct for this issue, the present data were baseline corrected at 50ms to minimize the interference of prior stimulation on baseline. One possible consideration would be to add pauses between syllables to correct the baseline issue. However, it may be that even very short pauses facilitate a different mechanism for speech segmentation (e.g., Buiatti, Peña, & Dehaene-Lambertz, 2008) by providing an additional acoustic cue beyond distributional information to the learner (e.g., Mueller, Bahlman, & Friederici, 2008). The second limitation is that using performance on an explicit measure, the 2AFC, to interpret implicit neural responses to implicitly learned material may not be an accurate reflection of what is learned during online segmentation. The 2AFC measure only allows a fairly coarse grouping based on accuracy scores, and other approaches may allow a more fine-grained approach to examining individual differences in learning. Some recent examinations of statistical learning have adopted implicit behavioural measures (e.g., Batterink et al. 2105) and more sensitive explicit measures (e.g., Siegelman et al. 2016) which may provide a better estimation of the outcome of statistical learning. In future research, the combination of these behavioural measures with online neuroimaging approaches (e.g., Batterink & Paller, 2017) would better elucidate the process of statistical learning.

4.4.1 Conclusions

Using a novel neuroimaging approach, the present study assessed the process of statistical learning in real time and uncovered a potential ERP marker that is sensitive to the statistical relationships between syllables in a word segmentation paradigm. Participants were exposed to an artificial language for 21-minutes, and ERPs were measured continuously throughout language exposure. Participants were grouped into high and low learners. High-learners showed sensitivity to the conditional structure of the speech stream, as indexed by an increased P200 to predictable, word-final syllables by the second minute of artificial language exposure. The ERPs from the low-learner group, however, did not seem to be modulated by increased language exposure. This finding fits with previous descriptions of the P200, which describe it as a marker of attentional allocation to structured material (e.g., Snyder et al., 2006). The results also have important implications for theories of statistical learning suggesting that those who demonstrate greater recognition of newly segmented words, that is, the high-learner group, show predictable neural responses to learned regularities within the speech stream.

4.5 References

Abla, D., Katahira, K., & Okanoya, K. (2008). On-line assessment of statistical learning by event-related potentials. *Journal of Cognitive Neuroscience*, *20*(6), 952-964.

Allison, T., Wood, C. C., McCarthy, G., (1986). The central nervous system. In: Coles, M.G.H., Donchin, E., Porges, S. (Eds.), *Psychophysiology: Systems, Processes, and Applications*. New York, NY: Guilford Press, pp. 5–25.

Batterink, L. J., Reber, P. J., & Paller, K. A. (2015). Functional differences between statistical learning with and without explicit training. *Learning & Memory*, *22*(11), 544-556.

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language*, *83*, 62-78.

Batterink, L., & Neville, H. (2013). Implicit and explicit second language training recruit common neural mechanisms for syntactic processing. *Journal of Cognitive Neuroscience*, *25*(6), 936-951.

Boh, B., Herholz, S. C., Lappe, C., & Pantev, C. (2011). Processing of complex auditory patterns in musicians and nonmusicians. *PLoS One*, *6*(7), e21458.

Buiatti, M., Peña, M., & Dehaene-Lambertz, G. (2009). Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses. *Neuroimage*, *44*(2), 509-519.

Craik, F. I., & Byrd, M. (1982). Aging and cognitive deficits. In *Aging and Cognitive Processes* (pp. 191-211). Springer, Boston, MA.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671-684.

Cunillera, T., Càmara, E., Toro, J. M., Marco-Pallares, J., Sebastián-Galles, N., Ortiz, H., ... & Rodríguez-Fornells, A. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *Neuroimage*, *48*(3), 541-553.

Cunillera, T., Toro, J. M., Sebastián-Gallés, N., & Rodríguez-Fornells, A. (2006). The effects of stress and statistical cues on continuous speech segmentation: an event-related brain potential study. *Brain Research*, *1123*(1), 168-178.

De Diego-Balaguer, R., Toro, J. M., Rodríguez-Fornells, A., & Bachoud-Lévi, A. C. (2007). Different neurophysiological mechanisms underlying word and rule extraction from speech. *PLoS One*, *2*(11), e1175.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9-21.

Francois, C., & Schön, D. (2011). Musical expertise boosts implicit learning of both musical and linguistic structures. *Cerebral Cortex*, *21*(10), 2357-2365.

François, C., Jaillet, F., Takerkart, S., & Schön, D. (2014). Faster sound stream segmentation in musicians than in nonmusicians. *PloS One*, *9*(7), e101340.

French, B., & Cottrell, G. (2014, January). TRACX 2.0: A memory-based, biologically-plausible model of sequence segmentation and chunk extraction. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 36, No. 36).

French, R. M., Addyman, C., & Mareschal, D. (2011). TRACX: A recognition-based connectionist framework for sequence segmentation and chunk extraction. *Psychological Review*, *118*(4), 614.

George, E. M., & Coch, D. (2011). Music training and working memory: an ERP study. *Neuropsychologia*, *49*(5), 1083-1094.
Godey, B., Schwartz, D., De Graaf, J. B., Chauvel, P., & Liegeois-Chauvel, C. (2001). Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. *Clinical Neurophysiology*, *112*(10), 1850-1859.
Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*(5), 431-436.
Graf-Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science, 18,* 254-260.
Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*(4108), 177-180.
Hyvärinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural Networks*, *13*(4-5), 411-430.
Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621-647.
Lopez-Barroso, D., de Diego-Balaguer, R., Cunillera, T., Camara, E., Münte, T. F., & Rodríguez-Fornells, A. (2011). Language learning under working memory constraints correlates with microstructural differences in the ventral language pathway. *Cerebral Cortex*, *21*(12), 2742-2750.
Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 213.
Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), B101-B111.
Mirman, D. (2016). *Growth Curve Analysis and Visualization using R*. CRC Press.
Mueller, J. L., Bahlmann, J., & Friederici, A. D. (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.
Paller, K. A., & Kutas, M. (1992). Brain potentials during memory retrieval provide neurophysiological support for the distinction between conscious recollection and priming. *Journal of Cognitive Neuroscience*, *4*(4), 375-392.
Palmer, S. D., & Mattys, S. L. (2016). Speech segmentation by statistical learning is supported by domain-general processes within working memory. *The Quarterly Journal of Experimental Psychology*, *69*(12), 2390-2401.
Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science, 298*(5593), 604-607.
Perruchet, P., & Vinter, A. (1998). PARSER: A model for word segmentation. *Journal of Memory and Language*, *39*(2), 246-263.
Pioneer Hill Software. (2008). *Spectraplus* (Version 5.0.26.0). Poulsbo, Washington: Pioneer Hill.
Polich, J. (2003). Theoretical overview of P3a and P3b. In *Detection of change* (pp. 83-98). Springer, Boston, MA.
Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128-2148.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*(6), 855-863.
Reinke, K. S., He, Y., Wang, C., & Alain, C. (2003). Perceptual learning modulates sensory evoked response during vowel segregation. *Cognitive Brain Research*, *17*(3), 781-791.
Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in cognitive sciences*, *11*(6), 251-257.
Saffran, J. R. (2001). Words in a sea of sounds: The output of infant statistical learning. *Cognition*, *81*(2), 149-169.
Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926-1928.
Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*(4), 606-621.
Sanders, L. D., & Neville, H. J. (2003). An ERP study of continuous speech processing: I. Segmentation, semantics, and syntax in native speakers. *Cognitive Brain Research*, *15*(3), 228-240.
Sanders, L. D., Newport, E. L., & Neville, H. J. (2002). Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. *Nature Neuroscience*, *5*(7), 700.
Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime 1.0. Pittsburgh, PA: Psychological Software Tools.
Siegelman, N., Boegarts, L., & Frost R. (2016). Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behavioural Research Methods*, 1-15.
Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Philosophical Transactions of the Royal Society of London B*, *372*(1711), 20160059.
Snyder, J. S., Alain, C., & Picton, T. W. (2006). Effects of attention on neuroelectric correlates of auditory stream segregation. *Journal of Cognitive Neuroscience*, *18*(1), 1-13.
Sound Forge Audio Studio (2013). *Sony Creative Software Inc*.
SpectraPlus software (2008). *Pioneer Hill Software.*
Spreng, M. (1980). Influence of impulsive and fluctuating noise upon physiological excitations and short-time readaptation. *Scandinavian audiology. Supplementum*, (Suppl 12), 299.
Team, R. C. (2016). R: a language and environment for statistical computing. V. 3.3. 1. Vienna: R Foundation for Statistical Computing.
Thiessen, E. D., & Erickson, L. C. (2013). Beyond word segmentation: A two-process account of statistical learning. *Current Directions in Psychological Science*, *22*(3), 239-243.
Thiessen, E. D., & Pavlik Jr, P. I. (2013). iMinerva: a mathematical model of distributional statistical learning. *Cognitive Science*, *37*(2), 310-343.
Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: use of stress and statistical cues to word boundaries by 7-to 9-month-old infants. *Developmental psychology*, *39*(4), 706.
Triesman, A. M. (1964). Selective attention in man. *British Medical Bulletin*, *20*(1), 12-16.

Vasuki, M., Pragati, R., Sharma, M., Ibrahim, R. K., & Arciuli, J. (2017). Musicians' online performance during auditory and visual statistical learning tasks. *Frontiers in Human Neuroscience*, *11*, 114.

Zuijen, T. L. V., Sussman, E., Winkler, I., Näätänen, R., & Tervaniemi, M. (2004). Grouping of sequential sounds—an event-related potential study comparing musicians and nonmusicians. *Journal of Cognitive Neuroscience*, *16*(2), 331-338.

Chapter 5: General Discussion

Despite the growing research on statistical language learning, much of the prior research has only focused on the outcome of lab-based statistical learning paradigms, and few studies have examined the underlying processes involved in statistical learning. In this thesis, I addressed this issue by combing behavioural, clinical, and neuroimaging approaches to examine both the cognitive and neural processes involved in statistical learning. There were two main aims of this combined approach. First, I wanted to expand on the existing research examining how domain-general and domain-specific cognitive skills contribute to statistical learning. Second, I wanted to develop a novel, implicit measurement approach to examine the process of statistical language learning in real-time. In this chapter, I will begin by summarizing the main findings from Chapters 2, 3, and 4 addressing these aims. Next, I will make recommendations for future research based on these findings. I will then close with some comments on how statistical learning fits into our understanding of language acquisition more broadly.

5.1 Relevant Findings

5.1.1 Domain-general and domain-specific contributions to statistical language learning

Earlier theories viewed statistical learning as a domain-general process (Saffran & Thiessen, 2007), wherein similar computations were made across stimuli from different domains. For instance, the transitional probabilities between auditory linguistic and musical sequences were thought to be processed in much the same way (e.g., Saffran, Johnson, Aslin, & Newport, 1999). However, subsequent research demonstrated both qualitative and quantitative differences in processing statistical regularities across different domains (Conway & Christiansen, 2005). More recently, a theoretical account described how statistical learning across domains may be differentiated by distinct processing constraints across modalities. Domain-generality, on the other hand, arises from similar domain-general computations across domains (Frost, Armstrong, Siegelman, & Christiansen, 2015). Further still, other accounts (e.g., Arciuli, 2017; Palmer & Mattys, 2016) have suggested that statistical learning may be supported by distributed domain-general skills, including working memory or attention. I was interested in further examining the domain-specific and domain-general processing constraints on statistical learning. To do this, I examined how processing limitations may be instantiated behaviourally, or manifest as a result of a result of a processing impairment associated with disproportionate deficits in language learning, known as developmental language disorder (DLD).

The study reported in Chapter 2 provided evidence that domain-general working memory interference disrupted statistical language learning. In this study, interference effects were compared across tasks imposing a low or high working memory demand, and that were verbal or visuospatial in nature. It was assumed if working memory resources are involved in statistical learning, any interference effects on statistical learning would be due to the working memory demands of the secondary task. Analyses revealed that statistical learning was impaired for those engaged in a high-demand working memory task across both verbal and visuospatial task conditions. Interestingly, engagement in either of the low-demand working memory tasks did not interfere with statistical learning. This finding suggested that working memory may be involved in statistical learning in a domain-general way, which was consistent with earlier reports (e.g., Palmer & Mattys, 2016).

Overall, the findings from this study have important implications relevant to theories of statistical learning. First, theoretical accounts which suggest that statistical learning involves shared and distributed component processes, including working memory, are consistent with the findings in Chapter 2 (e.g., Arciuli, 2017). Specifically, it may be that the domain-general interference effects were due to the engagement of partially shared memory resources modulating the encoding of both the auditory language sequence and the working memory task stimuli. Memory-based accounts of statistical learning suggest that statistical learning arises from a set of memory processes involved in the extraction, encoding, and integration of statistical regularities (e.g., Thiessen & Pavlik, 2013; *see* Thiessen, 2017). It is likely that the constant updating and monitoring involved in the working memory task would have limited the memory resources required to successfully segment the language stream. Taken together, these results suggest that there is some involvement of domain-general working memory in statistical learning, which could be more fully examined by investigating working memory involvement in non-linguistic statistical learning tasks.

The study reported in Chapter 3 is an extension of the findings reported in Chapter 2 in examining how domain-general or domain-specific skills are associated with statistical learning across domains. In Chapter 3, I was interested in whether cognitive skills, including language and working memory, were associated with statistical learning abilities in school-aged children. To examine statistical learning across domains, participants completed both a statistical language learning task and a visual statistical learning task. Importantly, this study included children with developmental language disorder (DLD). Children with DLD have a relatively greater impairment in language, and examining their performance on verbal and visuospatial statistical learning tasks allowed me to uncover whether language deficits are associated with differential statistical learning outcomes across domains. The main finding from this study was that children with a language impairment were unimpaired on the statistical language learning task. In fact, they performed above chance while typically developing children did not differ from chance-level performance. However, it is important to note that the two groups did not statistically differ in their performance on this task. This finding was surprising, as it was inconsistent with previous meta-analytic findings showing statistical word segmentation deficits for children with DLD (e.g., Lammertink, Boersma, Wijnen, & Rispens, 2017), and failed to replicate a study that was methodologically similar to the study reported in Chapter 2 (e.g., Evans, Saffran, & Robe-Torres, 2009). The second main finding was that both the DLD and typically developing groups failed to show learning in the visual statistical learning task. This finding is somewhat difficult to interpret, however. It may simply be that this task was unsuitable for a developmental population. In light of this, the findings from the visual statistical learning task are not be useful for informing hypotheses regarding the domain-specificity or domain-generality of a statistical learning impairment in those with DLD. Finally, there was no evidence to support an association between statistical learning in either domain and any of the cognitive or linguistic measures assessed in this sample.

One clear implication from these findings is that a relatively domain-specific language impairment was not associated with poor statistical language learning. It is possible that children with DLD are unimpaired on statistical learning tasks, which calls into question previous findings demonstrating a statistical learning deficit for this group (e.g., Evans et al., 2009; Mainela-Arnold & Evans, 2014; Lammertink et al., 2017; Lum, Conti-Ramsden, Morgan, & Ullman, 2014; Obeid, Brooks, Brooks, Powers, Gillespie-

Lynch, & Lum, 2016). However, it is important to consider that the performance on the statistical language learning task was generally meager, with scores only marginally above chance for the DLD group. One finding that is clear is that statistical learning across both the verbal and visuospatial task variants was not associated with performance on any of the language or cognitive measures. This finding is perhaps not surprising as previous studies have reported no association between verbal and visuospatial statistical learning (Siegelman & Frost, 2015). Additionally, the lack of a correlation between statistical learning tasks and other aspects of cognition assessed in Chapter 3 may be due to the nature of these tasks. Statistical learning is largely viewed as an implicit task, while the cognitive skills measured, including working memory, language, and intelligence, all involve explicit processing (*see* Arciuli, 2017). Given this, it may be that these distinct implicit and explicit processing tasks are unrelated, or that implicit statistical learning does not draw on these broader, explicit abilities. There is a continued interest in examining how statistical learning relates to other cognitive processes, and future work will need to be innovative in how it assesses the way in which these components are involved in statistical learning.

The results from Chapter 3 are rather inconclusive in uncovering domain-specific and domain-general constraints on statistical learning in a clinical population. However, I suspected that the two-alternative forced-choice (2AFC) test was inaccurate in capturing statistical learning skills in the school-aged sample. It was worthwhile considering some of the limitations of this test, and exploring new methodological approaches that accurately elucidate the process of statistical learning. I will now turn to how I addressed this issue directly in Chapter 4.

5.1.2 Developing a novel measure to assess statistical learning

This thesis was primarily concerned with uncovering the process of statistical learning. Yet, this venture was inhibited from the outset as the commonly employed outcome measure for statistical learning, specifically the 2AFC task, may not be not well-suited for this goal. It is my belief that the 2AFC outcome measure used in Chapter 2 and Chapter 3 underestimated total statistical learning for both adults and children, which is consistent with recent findings demonstrating that the explicit 2AFC test underestimates implicit statistical learning (e.g., Batterink, Reber, Neville, & Paller, 2015). To overcome this, the study presented in Chapter 4 investigated a novel approach to measure statistical learning in real-time using electroencephalography (EEG). In this study, event-related potentials (ERPs) were measured in response to predictive, word-final syllables during exposure to a structured, artificial language. Specifically, changes in P200 amplitude over the exposure period were compared across "low" and "high" statistical learner groups. I found that for "high" learners, the P200 in response to word-final syllables was maximal at the 2nd minute of language exposure, and that the amplitude of this component attenuated with subsequent exposure. The notion that the high-learner group may shift their segmentation strategy, or may shift their attention to additional segmentation cues, is consistent with behavioural work showing that listeners shift their strategy from tracking words to uncovering the underlying structure when the signal contains cues that may facilitate this process (Peña, Buatti, Nespor, & Mehler, 2011). This pattern was qualitatively different from the "low" learner group, who showed a gradual linear increase in the P200 response over the first 4 minutes of language exposure.

The results of Chapter 4 are important for a number of reasons. First, they demonstrated that ERPs measured in real-time in response to structured linguistic material can be used to examine qualitative differences in statistical learning. This is a methodological contribution as it demonstrates sensitivity to distributional regularities using an indirect and implicit measure that captures the process of learning, not just the outcome. Second, the measurement of ERPs can easily be adapted to different developmental groups. The demonstration of reliable ERP responses following statistical learning in adults provides the groundwork for further exploration in other age groups, including infants (e.g., Teinonen, Fellman, Näätänen, Alku, & Huotilainen, 2009), and children. Finally, the findings from Chapter 4 have significant implications for statistical learning theories. These results indicate "high" learners are shifting their focus quickly and briefly to those structural regularities that are meaningful in the speech stream. The increased P200 response to word-final syllables in the first two minutes of language exposure (for the "high" learner group) may be indicative of the rapid encoding the regularities between syllables (e.g., de Diego-Balaguer, Toro, Rodríguez-Fornells, & Bachoud-Lévi, 2007). This has clear applicability to statistical learning theories that suggest that listeners must not only detect the regularities within the input, but encode these segmented units within memory (e.g., Perruchet & Vinter, 1996; Thiessen & Pavlik, 2013). Further work may be needed to uncover the neural indices of this elaborative processing in statistical word segmentation, and add additional empirical support for these memory-based accounts.

5.2 Recommendations and Directions for Future Research

The findings from this thesis add to the existing literature supporting the involvement of domain-general processes in statistical learning, and how statistical learning can be measured in real-time. However, the present findings also generate some new questions that motivate future work. The issue of developing appropriate measures to uncover the process of statistical learning was directly addressed in this thesis, however, future research should assess both the replicability of these findings and the applicability of these findings to theoretical accounts of statistical learning. Moreover, the findings reported in Chapter 4 of this thesis are necessarily constrained in that the comparisons of the ERPs were made across groups dichotomized on a potentially inaccurate explicit outcome measure (e.g., Siegelman, Bogaerts, & Frost, 2016). An appropriate application of these findings would be to uncover a continuous statistical learning measure that can account for the electrophysiological differences across individuals. Additionally, future investigations could examine whether the real-time ERP approach used in this study is in fact a valid measurement of statistical learning by corroborating these findings with other implicit measures of statistical learning (e.g., Batterink et al., 2015).

An additional area for future research would be to develop a methodology to examine the association between explicit cognitive abilities, such as working memory, and implicit statistical learning. As mentioned previously, it is difficult to align these disparate skills experimentally, and there is a need for innovative research techniques to examine these relationships. One notion that has not been fully considered is how implicit working memory is related to statistical learning. Implicit working memory has not been widely studied, but relatively recent research has suggested that at least some aspects of working memory can operate outside of conscious awareness, and can be recruited without conscious attention (Hassin, Bargh, Engell, & McCulloch, 2009). Some researchers have suggested implicit working memory may be involved in statistical learning (e.g., Arciuli & Simpson, 2011; Janacsek & Nemeth, 2013), although this possibility has not been examined. To test this prediction, we would first need to provide a formal and testable model incorporating the distributed processes involved in statistical learning, including processes such as implicit working memory and other cognitive skills, and the relationship between these processes.

It is important for future research to consider multiple dimensions of language learning simultaneously in statistical learning paradigms. It has been acknowledged that there are multiple interacting layers of language that are instrumental to language acquisition (e.g., Saffran, 2014); For instance, the interaction of sounds and meanings in the process of word learning (*see* Graf Estes, Evans, Alibali, & Saffran, 2007). Previous work has shown that multiple statistical cues do interact during language learning (e.g., Babineau & Shi, 2014; Cunillera, Toro, Sebastian-Galles, & Rodríguez-Fornells, 2006; Endress & Hauser, 2010; Finn & Hudson Kam, 2008; Thiessen & Saffran, 2003). One possibility that arises, given the findings from Chapter 4, is that there is a dynamic nature to learners' attention to statistical cues during word segmentation. This raises the possibility that other linguistic cues may interact with word segmentation cues in an iterative or combinatorial fashion. It would be valuable if future research continued this exploration into how multiple sources of statistical information interact dynamically during language acquisition, particularly at early developmental time points.

5.3 A Role for Statistical Learning in Language Acquisition

An important note on the studies presented in this thesis, and many other studies of statistical learning, is that lab-based statistical learning paradigms are necessarily devoid of the many interacting cues that may help a language learner acquire their native language. However, these approaches are, in fact, advantageous for researchers as it allows us to isolate cues that may be meaningful for specific aspects of language learning. Indeed, one of the goals for the first studies on speech segmentation was to isolate the transitional probability relationships between syllables to demonstrate how these regularities uniquely guide word segmentation (e.g., Jusczyk & Aslin, 1995; Saffran, Aslin & Newport, 1996; Saffran, Newport, & Aslin, 1996). Following these initial studies using relatively sparse linguistic input, work from Pelluchi and colleagues (2009a; 2009b) and Hay, Pelucchi, Graf-Estes, and Saffran (2011) demonstrated that transitional probabilities can be applied to word segmentation in natural languages, suggesting that even in a noisier linguistic context, there is a role for distributional regularities between syllables. Additionally, in the present thesis, isolating the role of transitional probabilities in word segmentation allowed me to investigate both the role of other cognitive processes in word segmentation, and to measure statistical learning in real-time without contamination from confounding linguistic cues. Thus, there are clear and useful applications for statistical learning paradigms in the lab.

Although there are advantages to these statistical learning approaches, no one has ever claimed that transitional probabilities are the only route to word segmentation. It is important to acknowledge in any account of statistical learning that statistical cues are merely one of the many tools available in the language acquisition process. There are a myriad of other cues available to the language learner, including rhythm and pauses (e.g., Johnson & Jusczyk, 2001; Thiessen, Hill, & Saffran, 2005) and native-language stress patterns (e.g., Echols, Crowhurst & Childers, 1997; Thiessen & Saffran, 2007) that may help guide word segmentation. Extra-linguistic cues are also valuable to the language learner, including hearing words in isolation (e.g., Brent & Siskind, 2001; Fernald & Morikawa, 1993), referential statistics (e.g., Smith & Yu, 2008), and social interaction (Kuhl, Tsao & Liu, 2003). It is this complexity in our linguistic input that makes the study of language acquisition a rich area of research.

Infants are incredibly adept patterns learners (*see* Lewkowicz, Schmuckler, & Mangalindan, 2018) and likely take advantage of any combination of available cues to help them discern the complex linguistic environment that surrounds them. We have certainly come a long way from thinking of language acquisition as the manifestation of simple associative learning responses (e.g., Skinner, 1957), or the result of a specialized language acquisition device (e.g., Chomsky, 1965), and have gone on to consider the complex yet predictable sea of sounds (Saffran, 2001) surrounding the naïve language learner.

5.4 References

Arciuli, J. (2017). The multi-component nature of statistical learning. *Philosophical Transactions of the Royal Society of London B*, *372*(1711), 20160058.

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

Babineau, M., & Shi, R. (2014). Distributional cues and the onset bias in early word segmentation. *Developmental psychology*, *50*(12), 2666.

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language*, *83*, 62-78.

Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, *81*(2), B33-B44.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT press.

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

Cunillera, T., Toro, J. M., Sebastián-Gallés, N., & Rodríguez-Fornells, A. (2006). The effects of stress and statistical cues on continuous speech segmentation: An event-related brain potential study. *Brain Research*, *1123*(1), 168-178.

De Diego-Balaguer, R., Toro, J. M., Rodríguez-Fornells, A., & Bachoud-Lévi, A. C. (2007). Different neurophysiological mechanisms underlying word and rule extraction from speech. *PLoS One*, *2*(11), e1175.

Echols, C. H., Crowhurst, M. J., & Childers, J. B. (1997). The perception of rhythmic units in speech by infants and adults. *Journal of Memory and Language*.

Endress, A. D., & Hauser, M. D. (2010). Word segmentation with universal prosodic cues. *Cognitive Psychology*, *61*(2), 177-199.

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

Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development*, *64*(3), 637-656.

Finn, A. S., & Kam, C. L. H. (2008). The curse of knowledge: First language knowledge impairs adult learners' use of novel statistics for word segmentation. *Cognition*, *108*(2), 477-499.

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

Graf-Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science, 18,* 254-260.

Hassin, R. R., Bargh, J. A., Engell, A. D., & McCulloch, K. C. (2009). Implicit working memory. *Consciousness and Cognition*, *18*(3), 665-678.

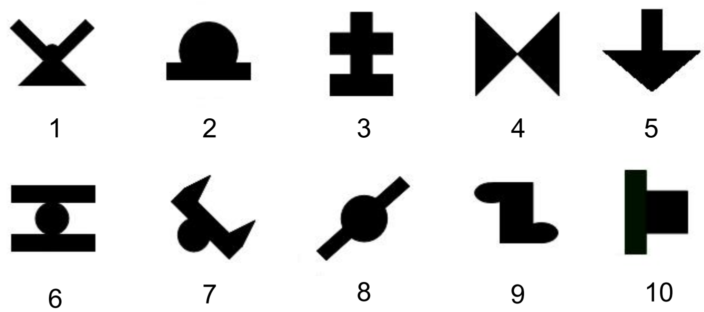
Hay, J. F., Pelucchi, B., Estes, K. G., & Saffran, J. R. (2011). Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology*, *63*(2), 93-106.

Janacsek, K., & Nemeth, D. (2013). Implicit sequence learning and working memory: correlated or complicated?. *Cortex*, *49*(8), 2001-2006.
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.
Jusczyk, P. W., & Aslin, R. N. (1995). Infants′ detection of the sound patterns of words in fluent speech. *Cognitive psychology*, *29*(1), 1-23.
Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, *100*(15), 9096-9101.
Lammertink, I., Boersma, P., Wijnen, F., & Rispens, J. (2017). Statistical Learning in Specific Language Impairment: A Meta-Analysis. *Journal of Speech, Language, and Hearing Research*, *60*(12), 3474-3486.
Lewkowicz, D. J., Schmuckler, M. A., & Mangalindan, D. M. (2018). Learning of hierarchical serial patterns emerges in infancy. *Developmental Psychobiology*, *60*(3), 243-255.
Lum, J. A., Conti-Ramsden, G., Morgan, A. T., & Ullman, M. T. (2014). Procedural learning deficits in specific language impairment (SLI): A meta-analysis of serial reaction time task performance. *Cortex*, *51*, 1-10.
Mainela-Arnold, E., & Evans, J. L. (2014). Do statistical segmentation abilities predict lexical-phonological and lexical-semantic abilities in children with and without SLI?. *Journal of Child Language*, *41*(2), 327-351.
Obeid, R., Brooks, P. J., Powers, K. L., Gillespie-Lynch, K., & Lum, J. A. (2016). Statistical learning in specific language impairment and autism spectrum disorder: A meta-analysis. *Frontiers in Psychology*, *7*, 1245.
Palmer, S. D., & Mattys, S. L. (2016). Speech segmentation by statistical learning is supported by domain-general processes within working memory. *The Quarterly Journal of Experimental Psychology, 69*(12), 2390-2401.
Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009a). Statistical learning in a natural language by 8-month-old infants. *Child Development*, *80*(3), 674-685.
Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009b). Learning in reverse: Eight-month-old infants track backward transitional probabilities. *Cognition*, *113*(2), 244-247.
Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science, 298*(5593), 604-607.
Perruchet, P., & Vinter, A. (1998). PARSER: A model for word segmentation. *Journal of Memory and Language*, *39*(2), 246-263.
Saffran, J. R. (2001). Words in a sea of sounds: The output of infant statistical learning. *Cognition*, *81*(2), 149-169.
Saffran, J. R. (2014). Sounds and meanings working together: Word learning as a collaborative effort. *Language Learning*, *64*(s2), 106-120.
Saffran, J. R., & Thiessen, E. D. (2007). Domain-general learning capacities. *Blackwell Handbook of Language Development*, 68-86.
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*(1), 27-52.

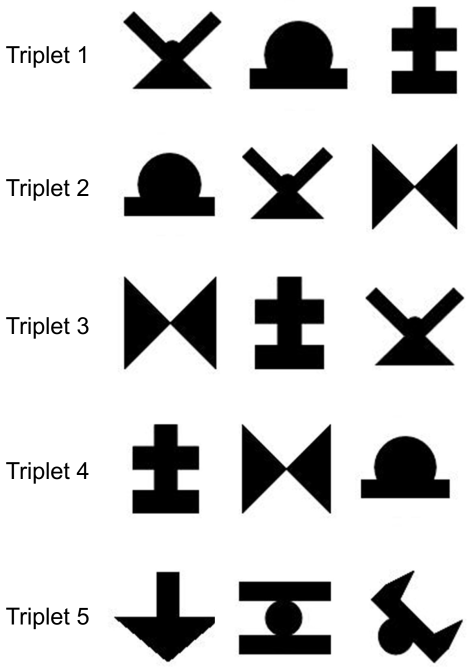
Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language, 35*, 606-621.
Siegelman, N., Boegarts, L., & Frost R. (2016). Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behavioural Research Methods*, 1-15.
Skinner, B. F. (1957). *Verbal behavior*. BF Skinner Foundation.
Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558-1568.
Teinonen, T., Fellman, V., Näätänen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neuroscience, 10*, 21.
Thiessen, E. D. (2017). What's statistical about learning? Insights from modelling statistical learning as a set of memory processes. *Philosophical Transactions of the Royal Society of London B*, *372*(1711), 20160056.
Thiessen, E. D., & Pavlik Jr, P. I. (2013). iMinerva: a mathematical model of distributional statistical learning. *Cognitive Science*, *37*(2), 310-343.
Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: use of stress and statistical cues to word boundaries by 7-to 9-month-old infants. *Developmental Psychology*, *39*(4), 706.
Thiessen, E. D., & Saffran, J. R. (2007). Learning to learn: Infants' acquisition of stress-based strategies for word segmentation. *Language Learning and Development*, *3*(1), 73-100.
Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, *7*(1), 53-71.

Appendices

Appendix 1: Shape inventory and training items for the visual statistical learning task



Note. Shapes are numbered 1-10, and correspond to the numbering used in Appendix 2.



Appendix 2: Summary of the 35 test items for the visual statistical learning task

Pattern Recognition

Question Type	*Item*	*Triplet*	*Target*			*Foil 1*			*Foil 2*			*Foil 3*		
Triplet-2AFC	1	1	1	2	3	6	5	3						
	2	2	2	1	4	1	3	7						
	3	3	4	3	1	9	7	8						
	4	4	5	6	7	6	5	3						
	5	5	8	9	10	4	1	5						
	6	1	1	2	3	8	1	4						
	7	2	2	1	4	1	3	7						
	8	3	4	3	1	7	2	3						
	9	4	5	6	7	8	1	4						
	10	5	8	9	10	9	10	5						
Triplet-4AFC	11	1	1	2	3	7	2	3	5	2	10	6	7	10
	12	2	2	1	4	9	10	5	9	7	8	4	1	5
	13	3	4	3	1	1	2	4	6	7	10	1	2	4
	14	4	5	6	7	5	2	10	7	2	3	9	10	5
	15	5	8	9	10	6	5	3	5	2	10	1	3	7
Pair-2AFC	16	1	2	3		1	6							
	17	2	1	4		4	8							
	18	3	3	1		10	9							
	19	4	6	7		7	8							
	20	5	8	9		7	4							
Pair-4AFC	21	1	1	2		5	1		7	1		2	10	
	22	2	2	1		8	5		2	5		3	2	
	23	3	4	3		3	8		5	7		8	4	
	24	4	5	6		4	1		7	4		9	3	
	25	5	8	9		5	7		9	1		10	5	

Pattern Completion

Question Type	*Item*	*Triplet*	*Sequence*			*Completion Choices*		
Triplet	26	1	1	?	3	**2**	1	9
	27	2	2	1	?	2	**4**	7
	28	3	?	3	1	10	5	**4**
	29	4	?	6	7	**5**	4	8
	30	5	8	?	10	3	**9**	6
Pair	31	1	1	?		**2**	10	1
	32	2	?	4		5	**1**	8
	33	3	4	?		2	9	**3**
	34	4	6	?		**7**	4	6
	35	5	?	9		7	**8**	3

Note. Shapes are numbered 1-10, and correspond to the shapes depicted in Appendix 1. Order of presentation was counterbalanced within each question type. For pattern completion items, the correct response amongst the completion choices is bolded.

Appendix 3: Correlations between statistical language learning performance and standardized measures

Measure	1	2	3	4	5	6	7	8
1. SLL	1.0	-.049	.010	.285	.109	.213	.227	.297
2. Age (Months)		1.0	.119	.460*	.295	.695**	.089	.225
3. CELF			1.0	.575*	.600*	.357	.545*	.691**
4. AWMA				1.0	.537*	.694**	.676**	.562
5. MAVA-Expressive					1.0	.609*	.676*	.306
6. MAVA-Receptive						1.0	.498*	.486*
7. WASI-Block Design							1.0	.605*
8. WASI-Matrix Reasoning								1.0

Note. * = $p < .05$; ** = $p < .001$, $n = 24$

Appendix 4: Correlations between visual statistical learning performance, statistical language learning performance, and standardized measures

Measure	1	2	3	4	5	6	7	8	9
1. VSL	1.0	-.159	.171	.247	.042	-.030	.032	.192	-.114
2. SLL		1.0	.158	-.067	.338	.009	.050	.095	-.109
3. Age (Months)			1.0	.067	.393	.209	.561*	.148	.089
4. CELF				1.0	.610*	.548*	.225	.681*	.525*
5. AWMA					1.0	.582*	.624*	.707**	.465
6. MAVA-Expressive						1.0	.580*	.297	.313
7. MAVA-Receptive							1.0	.410	.102
8. WASI-Block Design								1.0	.528*
9. WASI-Matrix Reasoning									1.0

Note. * = $p < .05$; ** = $p < .001$, $n = 19$; Note that the correlation values between the cognitive measures are different from those in Table 1. Only scores for those who completed the VSL task on follow-up testing are reported here, which included fewer children than the original sample.

Appendix 5: Ethics approval for the experiments reported in Chapters 2 and 4

Western

Department of Psychology The University of Western Ontario
Room 7418 Social Sciences Centre,
London, ON, Canada N6A 5C1
Telephone: (519) 661-2067Fax: (519) 661-3961

Use of Human Subjects - Ethics Approval Notice

Review Number	12 10 32	**Approval Date**	12 10 24
Principal Investigator	Lisa Archibald/Nicolette Noonan	**End Date**	13 10 23
Protocol Title	Working memory and learning		
Sponsor	n/a		

This is to notify you that The University of Western Ontario Department of Psychology Research Ethics Board (PREB) has granted expedited ethics approval to the above named research study on the date noted above.

The PREB is a sub-REB of The University of Western Ontario's Research Ethics Board for Non-Medical Research Involving Human Subjects (NMREB) which is organized and operates according to the Tri-Council Policy Statement and the applicable laws and regulations of Ontario. (See Office of Research Ethics web site: http://www.uwo.ca/research/ethics/)

This approval shall remain valid until end date noted above assuming timely and acceptable responses to the University's periodic requests for surveillance and monitoring information.

During the course of the research, no deviations from, or changes to, the protocol or consent form may be initiated without prior written approval from the PREB except when necessary to eliminate immediate hazards to the subject or when the change(s) involve only logistical or administrative aspects of the study (e.g. change of research assistant, telephone number etc). Subjects must receive a copy of the information/consent documentation.

Investigators must promptly also report to the PREB:
a) changes increasing the risk to the participant(s) and/or affecting significantly the conduct of the study;
b) all adverse and unexpected experiences or events that are both serious and unexpected;
c) new information that may adversely affect the safety of the subjects or the conduct of the study.

If these changes/adverse events require a change to the information/consent documentation, and/or recruitment advertisement, the newly revised information/consent documentation, and/or advertisement, must be submitted to the PREB for approval.

Members of the PREB who are named as investigators in research studies, or declare a conflict of interest, do not participate in discussion related to, nor vote on, such studies when they are presented to the PREB.

Clive Seligman Ph.D.

Chair, Psychology Expedited Research Ethics Board (PREB)

The other members of the 2012-2013 PREB are: Mike Atkinson (Introductory Psychology Coordinator), Rick Goffin, Riley Hinson Albert Katz (Department Chair), Steve Lupker, and TBA (Graduate Student Representative)

CC: UWO Office of Research Ethics

This is an official document. Please retain the original in your files

Appendix 6: Ethics approval for the experiment reported in Chapter 3

Western Research

Research Ethics

Western University Non-Medical Research Ethics Board
NMREB Amendment Approval Notice

Principal Investigator: Lisa Archibald
Department & Institution: Health Sciences\Communication Sciences & Disorders,Western University

NMREB File Number: 108053
Study Title: Language and Learning Abilities in Children
Sponsor: Natural Sciences and Engineering Research Council

NMREB Revision Approval Date: October 04, 2016
NMREB Expiry Date: August 18, 2017

Documents Approved and/or Received for Information:

Document Name	Comments	Version Date
Revised Western University Protocol	-Received Sept 29, 2016	

The Western University Non-Medical Science Research Ethics Board (NMREB) has reviewed and approved the amendment to the above named study, as of the NMREB Amendment Approval Date noted above.

NMREB approval for this study remains valid until the NMREB Expiry Date noted above, conditional to timely submission and acceptance of NMREB Continuing Ethics Review.

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario.

Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB.

The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000941.

Ethics Officer, on behalf of Dr. Riley Hinson, NMREB Chair

Western University, Research, Support Services Bldg., Rm. 5150
London, ON, Canada N6G 1G9 t. 519.661.3036 f. 519.850.2466 www.uwo.ca/research/ethics

Curriculum Vitae

NICOLETTE NOONAN, M.Sc.
The University of Western Ontario

Education

The University of Western Ontario 2014-Present
Ph.D. Psychology (*in progress*)
Advisors: Dr. Lisa Archibald & Dr. Marc Joanisse

The University of Western Ontario 2012-2014
M.Sc. Health and Rehabilitation Sciences
Speech and Language Sciences
Advisor: Dr. Lisa Archibald

The University of Western Ontario 2008-2012
B.A. Psychology, Honours Specialization
Graduated with distinction

Scholarships and Awards
Total funding: $136,900

Ontario Graduate Scholarship 2017-2018
Value: $15,000

Doctoral Research Excellence Award 2016-2017
Value: $5,000

Natural Science and Engineering Research Council PGS-D 2014-2017
Value: $63,000

Western Graduate Research Scholarship 2014-2018
Department of Psychology
Value: $30,000

Symposium on Research in Child Language Disorders, 2013
STAR student travel award
Value: $700

Western Graduate Research Scholarship 2012-2014
Department of Health Sciences
Value: $20,000

John A. McNee Award in Political Science 2009
Value: $1,200

The University of Western Ontario Scholarship of Excellence 2008-2009
Value: $2,000

Peer-Reviewed Publications

1. **Noonan, N. B.,** Redmond, S. M., & Archibald, L. M. D. (2014). Contributions of children's linguistic and working memory proficiencies to their judgments of grammaticality. *Journal of Speech, Language, and Hearing Research*.

Other Scholarly Work

1. Archibald L. M. D. & **Noonan N. B.** (2015). Processing deficits in children with language impairment. In E. Bavin & L. Naigles (Eds), *The Cambridge Handbook of Language, 2nd Edition.* United Kingdom: Cambridge University Press.

Peer-Reviewed Oral Presentations

1. Noonan, N. B., Joanisse, M. F. & Archibald, L. M. D. (2018) Examining the relationship between statistical learning and language abilities in school-aged children. Presented at the *Symposium on Research in Child Language Disorders*, Madison, WI. June 2018.

Peer-Reviewed Posters

1. Noonan, N. B., Archibald, L. M. D., & Joanisse, M. F. (2017) A window for word learning: Measuring dynamic neural responses during statistical language learning. Poster presented at the *Society for the Neurobiology of Language Annual Meeting*, Baltimore, MD. November 2017.
2. Noonan, N. B., Archibald, L. M. D., & Joanisse, M. F. (2017) ERPs index real-time extraction of statistical regularities in an artificial language speech stream. Poster presented at the *International Conference on Interdisciplinary Advances in Statistical Learning*, Bilbao, Spain. June 2017.
3. Noonan, N. B., Archibald, L. M. D., & Joanisse, M. F. (2017) Indexing moment-by-moment learning of statistical regularities using event-related potentials. Poster presented at the *Symposium on Research in Child Language Disorders*, Madison, WI. June 2017.
4. Noonan, N. B., Archibald, L. M. D., & Joanisse, M. F. (2016). Individual variation in the time course of statistical word segmentation: An ERP investigation. Poster presented at the *Symposium on Research in Child Language Disorders*, Madison, WI. June 2016.
5. Noonan, N. B., Archibald, L. M. D., & Joanisse, M. F. (2016). The output of statistical word segmentation: An ERP investigation. Poster presented at the *Canadian Society of Brain, Behaviour, and Cognitive Sciences*, Ottawa, ON. June 2016.
6. Noonan, N. B. & Archibald, L. M. D. (2015). The role of domain-specific processing resources in statistical language learning. Poster presented at *First Annual Statistical Learning Seminar*, San Sebastian, Spain. June 2015.
7. Noonan, N. B. & Archibald, L. M. D. (2015). Parallel implicit and explicit processing mechanisms in statistical language learning. Poster presented at *Symposium on Research in Child Language Disorders*, Madison, WI. June 2015.

8. Noonan, N. B. & Archibald, L. M. D. (2015). The role of multiple talkers in a statistical language learning paradigm. Poster presented at *Symposium on Research in Child Language Disorders*, Madison, WI. June 2015.
9. Noonan, N. B. & Archibald, L. M. D. (2014). Statistical language learning: Support from General Attention, or Verbal Capacity? Poster presented at the *Annual Meeting of the Psychonomic Society,* Long Beach, CA. November, 2014.
10. Noonan, N. B. & Archibald, L. M. D. (2014). Verbal implicit statistical learning: Evidence for dual-task interference. Poster presented at the *Canadian Society for Brain, Behaviour, and Cognitive Sciences 24th Annual Meeting.* Toronto, ON. July 2014.
11. Noonan, N. B. & Archibald, L. M. D. (2014). The relationship between implicit and explicit processing in statistical language learning. Poster accepted at the *Symposium on Research in Child Language Disorders*, Madison, WI. June 2014.
12. Noonan, N. B. & Archibald, L. M. D. (2013). Statistical learning of word boundaries in native and non-native speakers of English. Poster presented at *The British Psychological Society Developmental and Cognitive Sections Joint Conference*, Reading, United Kingdom. September 2013.
13. Noonan, N. B., Redmond, S. M., & Archibald, L. M. D. (2013). Differentiating linguistic and working memory demands in children's grammaticality judgments. Poster presented at the *Child Language Seminar*, Manchester, United Kingdom. June 2013.
14. Noonan, N. B., Redmond, S. M., & Archibald, L. M. D. (2013). Examining the effects of specific language impairment or co-occurring language and working memory impairment on children's grammaticality judgments. Poster presented at the *Symposium on Research in Child Language Disorders*, Madison, WI. June 2013

Non-Peer-Reviewed Posters

1. Noonan, N. B., Redmond, S. M., & Archibald, L. M. D. (2013). Differentiating linguistic and working memory demands in children's grammaticality judgments. Poster presented at the Western University Faculty of Health Sciences Research Day, London: Canada. February 2013.

Professional Experience

EEG and Sleep Lab Coordinator 2018-Present
BrainsCAN
The University of Western Ontario

Conference Co-Chair 2016-Present
Inspiring Young Women in STEM Conference

University Board of Governors Graduate Student Representative (elected at-large) *The University of Western Ontario*	2016-2018
University Governance and By-Laws Committee (appointed by the University Secretariat) *The University of Western Ontario*	2017-2018
Research Assistant to Dr. Ken McRae Department of Psychology *The University of Western Ontario*	2017
Research Assistant to Dr. Debra Jared Department of Psychology *The University of Western Ontario*	2017
Three Minute Thesis University Finalist *The University of Western Ontario*	2016
University Property and Finance Committee (appointed by the University Secretariat) *The University of Western Ontario*	2016-2018
Bishop Hellmuth Prize Committee Member (appointed by the VP Research) *The University of Western Ontario*	2016
Organizing Committee *Inspiring Young Women in STEM Conference*	2015-2016
University Disciplinary Action Committee, Reviewer (appointed by the University Secretariat) *The University of Western Ontario*	2015-2016
Internal reviewer for MClSc program in Driving Rehabilitation Therapy (appointed by the VP of Graduate and Postdoctoral Studies) *The University of Western Ontario*	2015
University Research Board, Graduate Representative (appointed by the University Senate) *The University of Western Ontario*	2014-2015
Research assistant to Dr. Lisa Archibald Language and Working Memory Lab, Department of Communication Sciences and Disorders *The University of Western Ontario*	2012-2014

Research student for Dr. Mary-Lou Vernon 2012
The University Laboratory School, Department of Psychology
The University of Western Ontario

Research assistant to Dr. Rod Martin 2011
The Humour Lab, Department of Psychology
The University of Western Ontario

Teaching Experience
(at *The University of Western Ontario*)

Language Development 2017
Psychology 3141F
Teaching Assistant/Guest Lecturer

History of Psychology 2017
Psychology 3950G
Teaching Assistant

Psychology of Language 2016
Psychology 2134A
Teaching Assistant/Guest Lecturer

Research Methods in Psychology 2014-2015
Psychology 2800E
Teaching Assistant

Healthology 2014
Health Sciences 2000B
Teaching Assistant

Editorial Experience
Ad-Hoc Reviewer:
Journal of Speech, Language, and Hearing Research
NeuroImage