

Electronic Thesis and Dissertation Repository

---

8-24-2021 10:00 AM

## The Role of Top-down Attention in Statistical Learning of Speech

Stacey Reyes, *The University of Western Ontario*

Supervisor: Batterink, Laura J, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Psychology

© Stacey Reyes 2021

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Cognition and Perception Commons](#), and the [Cognitive Psychology Commons](#)

---

### Recommended Citation

Reyes, Stacey, "The Role of Top-down Attention in Statistical Learning of Speech" (2021). *Electronic Thesis and Dissertation Repository*. 8090.

<https://ir.lib.uwo.ca/etd/8090>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact [wlsadmin@uwo.ca](mailto:wlsadmin@uwo.ca).

# THE ROLE OF ATTENTION IN STATISTICAL LEARNING

## Abstract

Statistical learning (SL) refers to the ability to extract regularities in the environment and has been well-documented to play a key role in speech segmentation and language acquisition. Whether SL is automatic or requires top-down attention is an unresolved question, with conflicting results in the literature. The current proposal tests whether SL can occur outside the focus of attention. Participants either focused towards, or diverted their attention away from an auditory speech stream made of repeating nonsense trisyllabic words. Divided-attention participants either performed a concurrent visual task or a language-related task during exposure to the nonsense speech stream, while control participants focused their attention to the speech stream. Visual attention was taxed through the classic Multiple Object Tracking paradigm, requiring tracking of multiple randomly moving dots. Linguistic attention was taxed through a self-paced reading task. Following speech exposure, SL was assessed with offline tests, including a post-exposure explicit familiarity rating task, and an implicit reaction-time (RT) based syllable detection task.

On the explicit familiarity rating measure, participants showed a reduction in learning when language-specific processing was taxed as compared to when visual resources were taxed. On the more implicit reaction time-based measure of SL, both divided-attention and full-attention controls performed comparably, all showing evidence of SL. These results suggest SL can proceed even when domain-specific (visual) resources are limited, but is compromised when more specific, language-related resources are taxed. These results offer insight into the neural cognitive underpinnings of SL and have exciting practical applications for improving adult second language acquisition.

*Keywords:* attention, statistical learning, implicit learning, explicit learning, language acquisition

## Summary for Lay Audience

Listening to an unfamiliar language can often be a disorienting experience. Natural speech is devoid of reliable pauses between words, making it difficult to determine where the words start and end. One way we can discover word boundaries is through statistical learning (SL), which refers to the ability to detect patterns in the world. SL is thought to play a key role in speech segmentation and language acquisition. Syllables within word boundaries tend to occur more frequently than syllables across word boundaries, and the ability to become sensitive to these statistical relationships between syllables is one way we can segment speech. Whether SL is automatic or requires focused attention is an unresolved question. The current study examined whether SL can occur outside the focus of attention, using both an explicit measure, as well as a more implicit reaction time-based task. Participants' either focused their attention towards, or diverted their attention away from an auditory nonsense speech stream. Participants who did not pay attention to the nonsense speech stream completed a task designed to tax either visual resources, or linguistic resources. Visual resources were taxed with an object tracking task, involving tracking a subset of randomly moving dots. Linguistic resources were taxed with a self-paced reading task, where participants read sentences and answered comprehension questions. Results showed that explicit learning was only reduced when linguistic resources were taxed, but unimpaired when visual resources were taxed. On the more implicit measure of SL, both divided-attention and full-attention controls performed comparably. These results demonstrate that SL can still occur to some extent when linguistic resources are taxed, and can occur uninterrupted with visual resources are taxed, providing support for the relative automaticity of SL. Practically, these results suggest language learners may engage in visual tasks when learning a new language.

# THE ROLE OF ATTENTION IN STATISTICAL LEARNING

## Acknowledgements

I would like to thank my supervisor LB for her patience, hard-work, encouragement, and for pushing me to my limit. Thank you for the long hours you spent on your thoughtful and precise edits. I would like to thank SVH, for his coding expertise and help with online data collection. SVH- you are the reason this experiment was made possible during the pandemic. Thank you to the CNLL lab for answering my tech/school questions, and for the grad school life talks that helped me feel less alone. Lastly, thank you JB for introducing me to language research.

# THE ROLE OF ATTENTION IN STATISTICAL LEARNING

## Table of Contents

Abstract.....	ii
Summary for Lay Audience.....	iii
Acknowledgements.....	iv
List of Figures.....	vii
List of Appendices.....	viii
Chapter 1.....	1
1 Introduction.....	1
1 Statistical learning.....	1
1.2.1 Mixed Findings in the Role of Attention in SL.....	3
1.2.2 Mixed Findings of the Role of Attention in Visual SL.....	7
1.3. Theoretical Rationale for a Possible Role of Top-down Attention in SL.....	11
1.4 Measures of SL.....	13
Chapter 2.....	17
2 Methods.....	17
2.1 Participants.....	17
2.2 Procedure.....	18
2.3 Stimuli and Experimental Tasks.....	22
2.3.2 SL Measures.....	25
2.4 Statistical Analyses.....	27
Chapter 3.....	29
3 Results.....	29
3.1 Rating Task.....	29
3.2.1 Target Detection.....	31
3.2.2 Target Detection- First Third of Experimental Task.....	32
3.2.3 Reaction time Prediction Effect.....	33
3.4 Accuracy on Concurrent Tasks (Divided-Attention Groups Only).....	34
3.5 Correlation Between Familiarity Rating and Target Detection Task.....	35

# THE ROLE OF ATTENTION IN STATISTICAL LEARNING

3.6 Post-task questionnaires.....	35
Chapter 4.....	35
4 Discussion.....	35
4.1 Attentional Impact when Phonological-Linguistic Resources are Taxed.....	36
4.2 Attentional Impact on Implicit versus Explicit Knowledge.....	39
4.3 Comparison to Past Findings.....	42
4.4 Limitations and Future Directions.....	44
4.5 Conclusions.....	47

List of Figures

*Figure 1.* Schematic of the study procedure.....20

*Figure 2.* Mean rating scores for words, partwords, and nonwords.....30

*Figure 3.* Mean reaction times (ms) for word-initial, word-medial, and word-final syllables.....32

*Figure 4.* Boxplot showing average RT percent decrease.....34

List of Appendices

Appendix A: Ethics Approval.....59

Appendix B: Consent Form.....60

Appendix C: Post-Task Questionnaires.....64

Appendix D: Multiple Object Tracking Sample.....66

Appendix E: Curriculum Vitae.....67



## Chapter 1

### 1 Introduction

#### 1.1 Statistical Learning

From the discovery of word boundaries in a foreign language to the categorization of novel objects, the ability to extract patterns in the environment plays a key role in human learning. This ability is known as *statistical learning (SL)*. SL plays a role in many different domains but has been especially noted for its role in language acquisition, and in particular, speech segmentation (Saffran et al., 1996). Natural speech consists of a continuous stream of sound, devoid of reliable acoustic cues to word units, making an unfamiliar language seem like an incomprehensible barrage of sound. By becoming sensitive to frequently occurring sound patterns in continuous speech, learners may be able to discover word boundaries in this barrage of sound. Despite the important theorized contribution of SL to language acquisition, the literature is relatively young, with research in this area beginning only 25 years ago, leaving many aspects of this phenomenon still to be elucidated. One major unresolved question concerns the role of top-down attention in statistical learning of speech. Given the vast array of concurrent environmental stimuli that is typically present along with the speech information, understanding the role of focused attention in SL will inform whether and how our limited capacity attentional system should be engaged to support SL.

*Statistical learning* first emerged as a key theoretical construct with the discovery that infants are sensitive to the relative probability at which syllables co-occur, referred to as their *transitional probability* (Saffran et al., 1996). The transitional probability between syllables XY is defined as the probability of Y given that X has occurred, and can be expressed by the following formula: frequency of XY/frequency of X. Transitional probabilities serve as a cue to word boundaries since syllables within words tend to have higher transitional probabilities, while

syllables across word boundaries have low transitional probabilities. For instance, to take the phrase “sleepy puppy”, the transitional probability between “slee” & “py” is higher than the transitional probability between “py” and “pu”, across the English language. In addition, the transitional probability between “pu” & “py” is higher than the transitional probability between “py” and “pu”. The relatively lower transitional probability between “py” and “pu” provides a word boundary cue.

In their seminal demonstration of SL, Saffran and colleagues (1996) exposed eight-month-olds to an artificial language comprised of four randomly repeating trisyllabic nonsense words. Words were concatenated (with the exception that no two words could repeat back to back, e.g. ABAB), and syllables were presented at a regular rate with constant intonation ensuring that the transitional probabilities between syllables served as the primary cue to word boundaries. For example, a sample presentation might be **pa-bi-ku-da-ro-pi-go-la-tu-ti-bu-do-daroti**... This arrangement ensures that the within-word transitional probability between adjacent syllables was 100% since a given word-initial syllable was always followed by the same word-medial syllable, which in turn was always followed by the same word-final syllable. For example, *pa* is always followed by *bi* which is always followed by *ku*. Between words, the transitional probability between one syllable and the next was 33%, since a given word-final syllable could be followed by the first syllable of any of the other three words in the language. For example, *ku* could have been followed by either *da*, *go*, or *ti*. After merely 2 minutes of exposure to the nonsense speech stream, in a separate test phase, infants showed longer listening times to foils (trisyllabic groupings never occurring in the nonsense language, e.g. *bi-da-ti*), compared to words from the language. This listening preference for novel items suggests that infants were able to segment words based on the statistical properties of the speech stream.

Importantly, these findings offer a solution to the long-standing speech segmentation problem: how are we able to perceive discrete lexical units in a continuous speech signal, which is devoid of reliable pauses between words? Saffran and colleagues' result suggests that gaining sensitivity to the statistical regularities in the speech signal may support the discovery of word boundaries.

Since this initial study in infants, SL has been demonstrated in children (Evans et al., 2009; Raviv & Arnon, 2017; Saffran et al., 1997), adults (e.g. Fiser & Aslin, 2001; Saffran et al., 1999) and animals (Hauser et al., 2001; Toro & Trobalon, 2005), as well as across other sensory modalities such as touch and vision (Hunt & Aslin, 2001; Fiser & Aslin, 2001; Turk-Browne et al., 2005). In older children and adults, rather than using a looking time measure, researchers typically administer a 2-alternative forced choice (2-AFC) recognition measure after the exposure phase. The 2-AFC requires participants to explicitly discriminate between words and foil items, with above-chance performance on this measure taken as evidence of SL. Perhaps one of the most interesting features of SL is that it occurs simply through passive exposure to the speech stream, in the absence of explicit instructions to uncover the underlying structure (e.g. Saffran et al., 1996a; 1996b; 1997). That SL can occur incidentally has introduced the question of whether SL requires attention.

### 1.2.1 Mixed Findings in the Role of Attention in SL- Behavioural studies

The role of attention in SL initially came into focus after researchers observed that SL could occur through passive exposure, without requiring effort or intention to learn. Using an incidental language learning paradigm, Saffran and colleagues (1997) were the first to introduce the notion of SL as an automatic process. In this pioneering study, children and adults completed a computer illustration cover task while an artificial language audiotope played in the

background. Importantly, participants were not informed of the linguistic nature of the audiotape, nor were they even told to listen to the language. Despite the lack of any instructions to attend to the audio stream, both children and adults showed above-chance performance on 2-AFC measures of SL, providing evidence of successful SL. Though attention was not explicitly manipulated, these findings indicate that SL can occur incidentally, outside the focus of direct attention.

On the other hand, research in infants has suggested attention to the speech stream may facilitate SL (Thiessen et al., 2005). Infants show evidence of SL to an artificial language spoken with intonation patterns characteristic of infant-directed speech, but not when the artificial speech stream was spoken with intonation patterns characteristic of adult-direct speech (Thiessen et al., 2005). Thiessen and colleagues (2005) posit that infant-directed speech may facilitate SL by promoting sustained attention to the speech.

Supporting the general notion that attention may facilitate SL, subsequent studies in adults that explicitly manipulated attention found compromised SL when attentional resources were taxed by a concurrent task. Toro et al. (2005) simultaneously presented participants with an artificial speech stream, along with three different dual tasks: (1) an auditory 2-back task using everyday noises such as a door slamming, (2) a visual 2-back task using everyday images such as furniture and body parts (presented at a rate of either 500ms, or 750 ms), and (3) a pitch detection task involving pitch changes in the artificial speech stream itself. SL via the 2-AFC measure was found to be significantly reduced among all divided-attention groups as compared to their full attention control counterparts. As well, all divided-attention groups performed at chance level on the 2-AFC, except for the group who completed the visual n-back distractor at the slower rate (750 ms), who still showed above chance performance despite having

significantly reduced performance compared to the passive listening group. This result is consistent with the well-established finding of a reduced attentional cost for concurrent tasks involving resources of a different domain, rather than of the same domain (e.g. Bayliss et al., 2003; Soto-Faraco & Spence, 2002; Treisman & Davies, 1973).

Similar findings were reported by Palmer and Mattys (2016), who found compromised SL when participants' phonological as well as visual resources were taxed. Participants performed either a 2-back rhyming task consisting of visually presented nonwords, or a 2-back visual task consisting of unnameable shapes. Using a 2-AFC measure of SL, both tasks were found to be equally disruptive to learning such that divided attention groups performed significantly worse than the full attention group, leading the authors to conclude SL is supported by domain-general attentional resources. However, both divided attention groups had above-chance performance on the 2-AFC, despite having significantly reduced SL performance as compared to the full attention groups, suggesting SL may still occur to some extent outside the focus of attention. Other work has painted a more nuanced picture of the role of attention in SL, reporting a graded effect of attention based on transitional probabilities. Using the same visual 2-back stimuli as Toro and colleagues (2005), Fernandes and colleagues found that reducing attention to the speech stream disrupted SL for words with low transitional probabilities (and thus less salient lexical cues), while SL continued to occur for high TP words (Fernandes et al., 2010). As well, low TP words under attentional load were not extracted above chance levels. These results demonstrate that under salient lexical cues in which the internal word TP is high, SL can continue to occur even when top-down attention is taxed.

Further evidence for the role of working memory in speech segmentation was shown by Lopez-Barroso and colleagues (2011), who found that preventing articulatory rehearsal impairs

SL, although learning was still above chance level. Although attention per se was not the prime area of interest, working memory and attention are generally known to be closely intertwined (e.g. Awh et al., 2006; Baddeley, 1974; Fougne, 2008); attention acts to select relevant information, which is then stored and maintained in an accessible state in working memory (e.g., Baddeley & Logie, 1999; Fougne, 2008). Lopez-Barroso and colleagues (2011) exposed participants to an artificial language and instructed them to utter the syllable “bla” continuously, in order to block the articulatory rehearsal subcomponent of the phonological loop in working memory. Learning (as assessed via the 2-AFC task) was significantly reduced for participants instructed to block articulatory rehearsal as compared to controls who did not receive these interference instructions. However, those instructed to block their articulatory rehearsal subcomponent still showed above-chance learning. These results suggest a supporting role for the phonological loop in statistically based speech segmentation. It is possible that disrupting a subvocalization strategy reduced the amount of time the information was active in working memory, making it more subject to decay.

In contrast to many of these findings, one recent study using an explicit familiarity measure, as well as an EEG measure and implicit reaction-time based measure of learning suggests that statistical learning can occur robustly even outside the focus of attention. Batterink and Paller (2019) manipulated participants’ attention with a visual 3-back task comprised of unnameable images while they listened to the nonsense speech stream. At test, participants completed a familiarity rating task in which they provided a 1-4 familiarity rating on words, part-words (2 syllables from one word and 1 syllable from another word) and non-words (all 3 syllables from 3 different words), as well as an implicit reaction-time based task that involves responding to target syllables within shortened segments of the speech stream. Learning on this

task is reflected by faster reaction times to syllables occurring later in the word as compared to earlier occurring syllables (e.g. Batterink et al., 2015; Batterink & Paller, 2017; Batterink & Paller, 2019). On the explicit rating task, divided-attention participants and full attention controls both showed similar, above-chance performance. In contrast, performance on the implicit reaction-time based measure was somewhat reduced when attention was divided. Although divided-attention participants had significantly slower reaction times overall as compared to controls, they still showed learning as indicated by the linear decrease in reaction times going from word-initial to word-final syllables. As well, the EEG measure, indexing SL during the exposure period, showed comparable word learning between the divided attention group, and full attention controls. These results demonstrate that taxing attention may somewhat impede long-term memory storage of the component words, as assessed by the implicit task. Nonetheless, given that both the EEG learning measure and explicit memory for the component words was not significantly different between the two attention groups, SL appears to occur robustly outside the focus of attention, even when participants are engaged in a highly demanding cover task.

### 1.2.2. Mixed Findings of the Role of Attention in Visual SL

As previously mentioned, SL has been shown to operate across other sensory modalities, suggesting that SL is not specific to language acquisition (e.g. Arciuli, 2017; Emberson et al., 2011; Siegelman & Frost, 2015). However, the degree to which SL is domain-general is still an open question, with recent work suggesting that SL may differ in important ways depending on the sensory modality and stimulus materials (e.g. Conway & Christiansen, 2005; 2006; Emberson et al., 2011; Fiser & Aslin, 2001). For instance, auditory SL performance is enhanced at faster presentation rates, while visual SL performance is better at slower rates (Emberson, 2011). More generally, audition is traditionally known to be equipped for temporal information,

while vision is known to be equipped for spatial information (Kubovy, 1988), limiting the generalization of auditory SL findings to visual SL. Nonetheless, to the extent that there are shared aspects of SL across different domains (Kirkham et al., 2002), studies on SL in other modalities such as vision may still provide insight into the role of attention in SL for speech sounds.

A typical visual SL paradigm involves presenting sequences of shapes, rather than syllables, that follow a temporal pattern. For example, Fiser and Aslin (2002) presented participants with a single shape that moved back and forth horizontally across the screen, behind a stationary occluding object, located in the middle of the screen. Each time the shape passed from behind the occluder, it would emerge as a different shape. This process would repeat, with a new shape emerging from the occluding with every horizontal pass. Unbeknown to the participant, there was an embedded shape triplet structure. Two-interval forced choice measures between a triplet and foil showed participants to have learned the structure.

Similar to studies in the auditory domain, research investigating the role of attention in visual SL has also produced mixed results. Turk-Browne and colleagues (2005) demonstrated an attentional effect to visual statistical learning. These researchers interleaved two temporal streams comprised of unnameable shape triplets, each stream with a distinct colour. Attention was manipulated by having participants complete a cover task in which they were to detect occasional shape repetitions in only one of the colours. On their forced-choice measure of learning, SL was found only in the attended stream, whereas learning was at chance level for the unattended stream. On their implicit reaction time measure of learning, reaction time effects only showed learning for the attended stream, and not for the unattended stream. These findings lead the authors to conclude that visual SL is “gated by selective attention” (Turk-Browne et al.,



2005). Emberson and colleagues (2011) found similar results. Using an adapted version of the interleaved design by Turk-Browne et al. (2005), Emberson and colleagues (2011) interleaved a stream of auditory triplets (nonsense syllables) with a stream of visual triplets (unnameable shapes), where triplets were statistically coherent. Participants were instructed to either attend to stimuli in the auditory modality, or the visual modality. Recognition measures showed decreased SL when attention was directed to the irrelevant modality, rather than when participants were tested on the modality of focus, suggesting a boosting effect of selective attention to SL. On the other hand, using a near-identical paradigm as the one used by Turk-Browne and colleagues (2005), Musz and colleagues (2014) found learning even for an unattended set. They found above chance performance on their explicit familiarity judgment measure for both attended and unattended triplets, with no significant difference in performance between attended vs unattended triplets. On their implicit reaction time measure, they found reaction times were significantly faster for the third position as compared to the first for both the unattended and attended streams, indicating similar learning. As these researchers directly replicated the paradigm by Turk-Brown et al., (2005), the discrepancy in results is not clear.

Using an adapted singleton paradigm, others have also shown visual SL to occur outside of attention. Duncan and Theeuwes (2020) investigated the question of whether visual statistical regularities can be learned even if they are not related to top-down goals. These researchers administered a training period designed to induce learning of the shape-location regularity. During this training period, participants were presented with a global array of shapes and performed a cover task where they were to determine whether the shapes were globally arranged in a circle, or a diamond. During these trials, a singleton distractor was present in a unique colour, located in a certain location more often than other locations. At test, participants were

presented with a standard additional singleton task requiring them to detect the one unique shape in the array (e.g., either a diamond among circles, or a circle among diamonds), while the unique colour distractor was randomly present in any of the location. The researchers reasoned that if participants learned the distractor location during the training phase, participants should show a suppression effect at test such that visual search would be faster when the shape singleton is located at the site where the colour distractor was most often located during the training phase. Indeed, this suppression effect was found, indicating participants learned the colour distractor location contingency during the training phase. As the colour distractor had nothing to do with the top-down goal of the learner (which was to determine the global shape), the researchers conclude visual SL occurs even when it is unrelated to the participant's top-down goals, suggesting that perhaps top-down attention is not needed for visual SL.

In summary, there are conflicting findings regarding the role of attention in SL. On the one hand, we see results from Batterink and Paller (2019), Duncan and Theeuwes (2020), and Musz and colleagues (2014), showing that SL can continue to occur outside the focus of top-down focused attention. On the other hand, findings from Emberson et al., (2011), Fernandes et al., (2010), Palmer and Mattys (2016), and Toro et al. (2005) show an effect of attention on SL, such that there is disrupted learning when top-down attention is taxed. The discrepancy in findings may be due to differences in sensitivity of the measures, as well as differences in attention manipulations. For example, for the one SL study by Batterink and Paller (2019) showing no attentional effect for explicit learning, one reason for the lack of an attentional effect may be the slow presentation rate of the distractor task. Though ERP indices of attention confirmed that the visual 3-back task captured attention, the presentation rate of the stimuli was relatively slow, with images appearing at a rate of 2.4-5.0 seconds (Batterink & Paller, 2019).

The slow presentation rate may have enabled participants to catch brief “snippets” of the nonsense language, enabling learning despite the concurrent task. Overall, the extent to which top-down attention is required for SL remains an unresolved question, and more work is needed using sensitive indices of SL, as well as additional attention-depleting distractor tasks.

### 1.3 Theoretical Rationale for a Possible Role of Top-down Attention in SL

Although SL has been found to occur without intention, this does not necessarily rule out the involvement of a higher-level, endogenous process was not involved. Firstly, since attention was not explicitly manipulated in studies demonstrating incidental learning, it is possible that passively listening to the speech stream (with no other concurrent stimuli) might have drawn participants to make sense of the stimuli. As Turk-Browne and colleagues (2005) point out, participants may wonder “how can I do nothing?”. In addition, previous studies have demonstrated a disruption to SL when working memory is taxed (e.g. Lopez-Barroso et al., 2011; Palmer & Mattys, 2016; Toro et al., 2005). Conway (2020) and Awh and colleagues (2006) point out that by definition, working memory and endogenous attention are closely related as they both involve holding select stimuli at a heightened state of availability. It is possible that at least some level of endogenous attention was directed at the speech stream, while still resulting in learning without intention.

The potential role of attention in SL may be understood through different models of theories of working memory, such as Baddeley & Hitch (1974). The highly influential Baddeley and Hitch (1974) three component model of working memory describes the central executive component as controlled by a limited capacity attentional system, directing the visuospatial-sketchpad and phonological loop to process select stimuli. In this model, the phonological loop is

specialized for the temporary storage and maintenance of verbal information (Baddeley 1986). Baddeley and colleagues (2003) propose that the phonological loop aids language learning by supporting the storage of unfamiliar sound patterns, which can then go on to form more long-term memory representations. Supporting this idea, Baddeley and colleagues (1998) have shown digit span performance to correlate with overall vocabulary in 4-13 year-old children. Further, longitudinal studies have demonstrated that nonword repetition ability at age 4 predicts vocabulary at age 5 (Gathercole & Baddeley, 1989). Corroborating these findings is a study on patient P.V., a native Italian speaker who is reported to have a circumscribed deficit to her phonological store. (Baddeley et al., 1988). P.V., showed comparable performance to healthy controls in learning novel word pairs in her native language, but was unable to learn the Russian equivalent of known Italian words, highlighting, the importance of the phonological loop in the acquisition of new words (Baddeley et al., 1988). Working memory may play a role in SL given that SL involves detecting regularities across a stream of speech elements, unfolding over time. A sufficient working memory span would be needed to hold enough of these elements to detect patterns across the speech stream.

Similarly, Janacsek & Nemeth (2013) speculate that working memory plays a role in implicit sequence learning, such that those with a greater working memory capacity have what they describe as a “larger window” to a sequence- that is, individuals are able to hold and maintain larger chunks of the sequence in working memory. Given that SL involves the tracking of statistical regularities among syllable groupings unfolding over time, it is possible that focused central attention may enable the storage of a larger number of trisyllabic chunks in working memory than one would normally be capable of holding without such focused attention. Other possible routes by which attention may benefit SL are offered by Palmer and Mattys (2016) who speculate that

central attention may enable the refreshing of statistically coherent syllable groupings in working memory, making them more likely to be stored in long-term memory. Another possibility is that attention may be needed to update the syllables held in working memory, removing syllable groupings that do not recurrently occur with one another (Palmer & Mattys, 2016).

#### 1.4 Measures of SL

As described previously, the most common approach to measuring SL is the 2-AFC recognition task. However, while the 2-AFC is a quick, simple, and frequently used task, it has been described to have a number of limitations (e.g. Arnon, 2020; Siegelman et al., 2017; Siegelman et al., 2017). SL involves implicit aspects of learning, yet participants are asked to make a conscious decision on the correct word (Christiansen, 2017). Additionally, performance on this task is not only sensitive to SL per se, but also to other peripheral abilities such as memory storage and retrieval, decision-making, and introspection abilities, all of which show substantial interindividual variability (Bors & Macleod, 1996; Siegelman et al., 2017; Unsworth, 2019). Furthermore, the task is argued to have high measurement error since a large proportion of participants perform at chance level (Siegelman et al., 2017). It has also been shown to have poor reliability across sessions, and low internal consistency (Arnon, 2020). In sum, the 2-AFC has a number of limitations. Although recognition tasks may be suitable as an explicit measure of SL, supplementary measures of learning may provide additional insights into SL, particularly to capture the more implicit aspects of learning.

As SL is thought to involve both implicit (knowledge that the individual is unable to consciously declare) and explicit memory traces (knowledge that the individual can consciously access) (e.g. Batterink et al., 2015), additional measures are needed to capture implicit

knowledge. Given that the 2-AFC largely taps into explicit memory, one step towards disentangling implicit and explicit memory contributions to SL is inclusion of the implicit reaction time-based target detection task (TDT) (e.g. Batterink, 2015; Franco et al., 2015; Turk-Browne et al., 2005). As described previously, in this task, participants listen to short segments of the artificial language, and make speeded responses to target syllables. Learning is reflected by faster reaction times to target syllables occurring in later, more predictable syllable positions in the word, as compared to word-initial targets. Batterink and colleagues (2015) have shown that this measure does not correlate with performance on the 2-AFC, and that a greater number of participants show learning on the TDT as compared to the 2-AFC. These results suggest the TDT is potentially a more sensitive measure of SL, and reveals more implicit word knowledge not captured by traditional recognition measures. The TDT used in conjunction with the 2-AFC may be used to reveal both implicit and explicit aspects of SL.

## 1.5 Current Study

The current study addresses the unresolved question of the role of top-down attention in SL. As prior work has mainly relied on the 2-AFC measure of learning, we used both an explicit task- familiarity rating task — and an implicit task- the TDT — to examine how different types of memory traces are affected by our attentional manipulations. In addition, we examined how two different types of attentional manipulations- calling upon linguistic versus visual resources- influence SL Both types of distractor tasks were designed to be highly demanding in order to maximally deplete attentional resources.

Participants were randomly assigned to one of four different groups, which required them to either complete a dual task while passively listening to the trisyllabic nonsense language, or to fully attend to the nonsense language (Figure 1). During exposure to the nonsense speech stream,

divided-attention participants either completed a concurrent reading task (taxing language-related resources), or a concurrent dot-tracking task (taxing visual-related resources) during exposure to the nonsense speech stream. Specifically, language-related resources were taxed with a *self-paced reading task*, in which participants read sentences one word at a time under time pressure, and answered comprehension questions. Visual attention was taxed with the classic *Multiple Object Tracking* (MOT) paradigm (Pylyshyn & Storm, 1988), in which participants tracked a subset of randomly moving dots on a screen, and indicate their positions. Two full-attention control conditions were also included, in which participants were presented with the same physical distractor stimuli of either the self-paced reading task or the MOT task, along with the nonsense speech stream, but instead were instructed to fully attend to the speech stream. The distractor stimuli were present in order to control for any effects of the physical stimuli on SL. Thus, there were a total of four groups- the divided attention group with the reading task, the divided attention group with the MOT task, and the two full attention control groups who were presented with the same physical distractor stimuli as their distractor counterparts.

After the exposure period, all four groups completed a familiarity rating task and the TDT. The rating task assesses explicit knowledge of the nonsense words, by presenting participants with either a trisyllabic word, part-word (a syllable pair from a language word + a syllable from a different word), or a non-word (three syllables which never occurred together in the language). Participants provide a 1-4 familiarity rating, allowing for a finer-grained measure of word knowledge (unlike the 2-AFC which requires participants to make a definitive choice), making it potentially a more sensitive measure as compared to the 2-AFC. The TDT was used as our implicit measure of learning, requiring participants to detect target syllables via keypress

upon hearing short segments of the speech stream. Learning is reflected by a RT prediction effect, or a linear decrease in RT as a function of syllable position. To the extent that the participant has learned the language, faster responses should occur for word final syllables that are predicted by both word-initial and word-medial syllables, while the slowest responses occur for initial syllables that have no predictive syllable cues, and intermediate RTs should occur for medial syllables which are predicted by the first syllable. This measure has been shown to reveal learning above and beyond what is captured in the explicit rating task (e.g. Batterink et al., 2015; Batterink and Paller, 2017).

Our goal was to answer three inter-related questions:

(1) *Is there an effect of top-down voluntary attention on SL?* On the familiarity rating task, we expect full-attention controls to show significantly better performance on the rating measure than divided-attention participants. On the TDT, we also expect full-attention controls to have a stronger reaction time prediction effect on the TDT, as indicated by a greater difference between word-initial and word-final reaction times.

(2) *Can SL occur in the absence of focused attention to the speech stream?* On the explicit measure, we would expect some learning to occur when across-domain, visual resources are taxed, but chance-level learning when within-domain, language-based resources are taxed, as the language-based distractor is expected to compete with the speech stream for neural resources. On our more implicit TDT measure of learning, we would expect learning to proceed even when attention is divided, given that implicit memory is known to be more resilient to interference (e.g. Jacoby & Dallas, 1981; Tulving et al., 1982).



(3) *Is there a differential impact based on the type of resources taxed (language-based vs visual)?* As alluded to above, we expect the within-domain, language-based self-paced reading task to exert the greatest impairment on learning as compared to the visual MOT distractor, as there would be greater competition for common neural resources.

## Chapter 2

### 2 Methods

#### 2.1 Participants

One hundred and eighty (45/group) native monolingual English-speaking adults aged 19-36 years (109 male) were recruited through Amazon Mturk paired with CloudResearch®, an online research study participant recruitment platform. Our final sample was comprised of 25 participants in the divided-attention visual group, 26 participants in the divided-attention reading group, 30 full-attention visual controls, and 30 full-attention reading controls. Participants were recruited from North America and were required to have a >95% approval rating on previously completed studies on the platform. Informed consent was obtained through both the CloudResearch® platform and through an electronic consent form via Qualtrics. A monetary compensation rate of \$10/hour of participation was provided to participants. Demographic data was collected through an electronic a pre-screening questionnaire assessing language history, handedness, neurological history, sex, and age. To participate in the study, participants were required to be native monolingual English speakers between the ages of 18-25, with no history of neurological disorders, with normal hearing, and normal or corrected-to-normal vision. Participants were randomly assigned to one of four groups: divided-attention visual (MOT task), divided-attention linguistic (self-paced reading task), full attention control group with MOT

stimuli present (visual controls), and the full attention control group with reading task stimuli present (linguistic controls).

### **Exclusion Criteria**

Participants had to pass the headphone check in which they were to indicate which of three tones was the lowest in volume, for a total of 6 trials. A minimum score of 5/6 was considered a pass. As well, participants had to perform the target detection measure with a total of <30 false alarms and <50 misses. This criteria was established after running a pilot study, showing that the majority of participants performed within this cut-off. Participants who did not fulfill all three (headphone check, <30 false alarms, and <50 misses) of these criteria were automatically excluded from the experiment. We excluded 39% of our sample, which is in line with previous online studies (Woods et al., 2015).

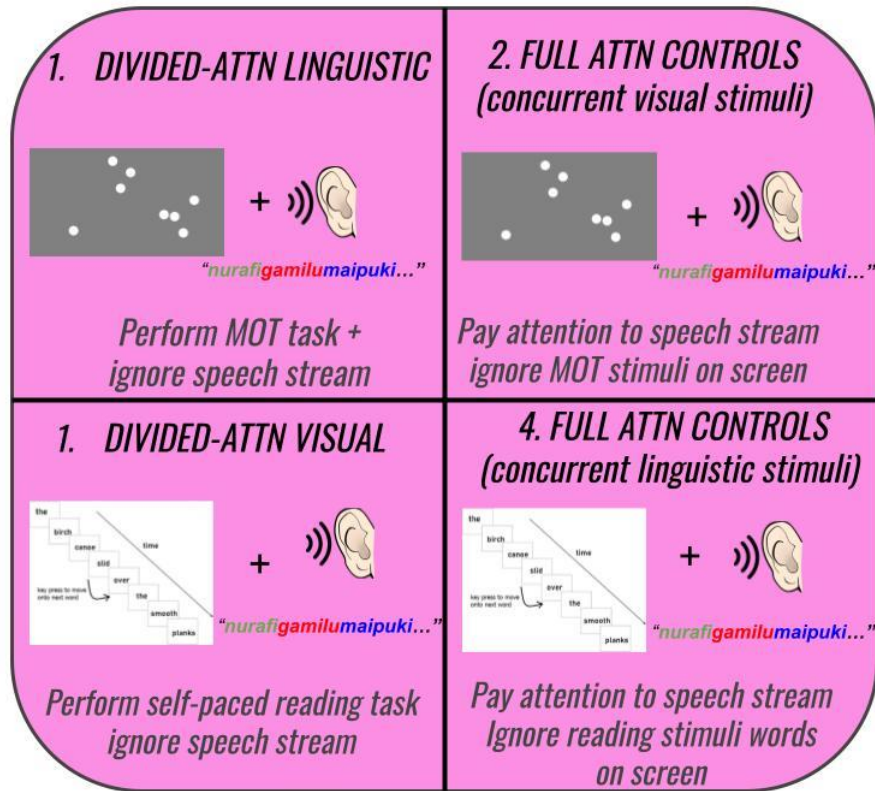
## **2.2 Procedure**

Informed consent was obtained via the Western University Qualtrics Online Survey platform. Once participants indicated their consent, they were then presented with a link at the end of the survey leading to the main experimental task hosted on Pavlovia, an online platform used to run behavioural science experiments. The experimental tasks were programmed on *jsPsych* (de Leeuw, 2015), a JavaScript library used in partnership with Pavlovia to create experiments.

The general structure of the online experiment is as follows: all participants first completed an audio quality check. This was followed by the exposure period in which they heard the nonsense language, and were simultaneously instructed to either complete a concurrent task (divided-attention groups), or focus on the sounds (control groups). Following the exposure period, participants answered the following personal, open-ended question as an attention check:

*What is one way your daily routine has changed during the pandemic?* This was then followed by the explicit rating task, which was followed by the implicit target detection task. Completing the TDT last minimized the influence of additional language learning on the explicit rating task, as participants are re-exposed to the nonsense language during the TDT. Lastly, participants completed a post-task questionnaire, allowing them to comment on the difficulty of the tasks, their awareness of the nonsense language, as well as any technical issues. Please see Figure 1 below for a schematic of the experimental paradigm.

1. Exposure period to nonsense speech stream (all 4 groups)



2. Measures of statistical learning

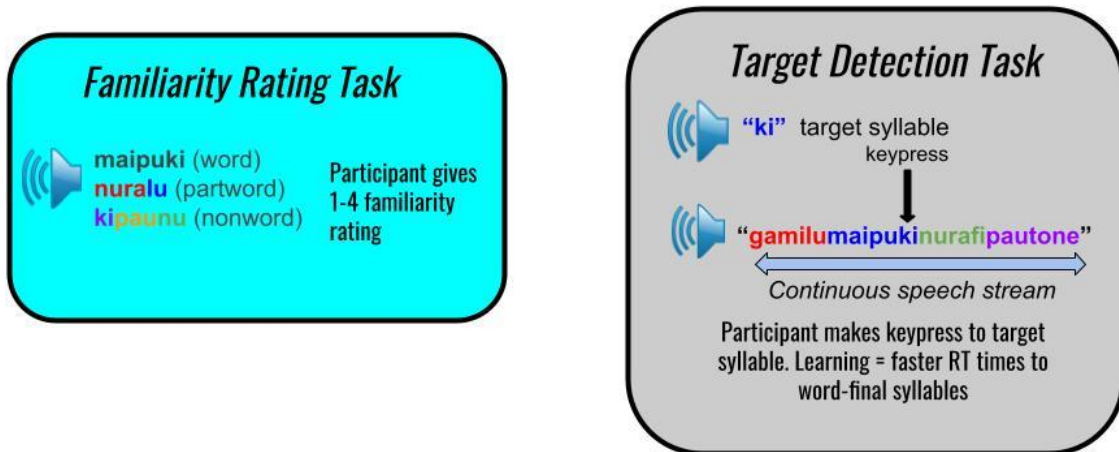


Figure 1. Schematic of the study procedure. All four groups completed a 6 minute exposure period, followed by behavioural measures of statistical learning.

Participants were exposed to the speech stream in twelve 30-second blocks for a total exposure time of 6 minutes. Divided attention participants simultaneously performed a concurrent task while listening to the speech stream (either tracking the motion of a subset of dots, or read sentences presented on the screen) during each 30-second block. Between each 30-second block, divided attention participants answered attention-check questions pertaining to the concurrent task under no time constraint. Full attention controls also simultaneously listened to the speech stream, while viewing the task stimulus presented to their divided attention counterparts in order to control for stimulus effects. However, they were instructed to pay attention to the sound and ignore the concurrent visual stimulus. As an attention check, occasional pauses occurred during the speech stream and controls were required to make a keypress as fast as possible in order for the speech stream to continue.

To control for the motor response required of divided attention participants during the question segment after each trial, control participants were instructed to type in a response of similar nature to their divided-attention counterparts. Divided-attention reading participants had to type in a one-word answer for the comprehension questions, whereas reading controls were presented with a random word chosen from the Harvard Sentence Inventory and were asked to type it in. Divided-attention visual participants were instructed to type in the number corresponding to the targets, whereas visual controls were presented with 2 or 3 random numbers and were asked to type them in.

Volume of the audio stream gradually increased at the beginning of each block, and gradually decreased at the end of each block to ensure that the speech stream did not abruptly begin with word-initial syllables, or end with word-final syllables, preventing participants from using the start and end of each auditory block as word boundary cues.

Although our online experimental format could not allow us to directly monitor participants' engagement, a number of attention checks were implemented to ensure participants were engaged as much as possible during the tasks. Full attention controls heard occasional pauses in the speech stream and were instructed to make a keypress as fast as possible in response to the cessation of sound. The pauses in sound presentation occurred for no longer than 2 seconds, and responses after 2 seconds were recorded as a null. Secondly, control participants were required to type in either random numbers or letters in between trials to control for the motor response required of their divided-attention counterparts who had to provide a response to the concurrent task. Though originally designed as control for motor activity, this also encouraged a form of active participation, and allowed us to identify any participants who were clearly not following task instructions. Aside from the very rare occurrence of a spelling mistake, all control participants performed this accurately. Third, at the end of the exposure period, all participants were asked an open-ended question: "what did you do as a leisure activity during the pandemic?" This allowed us to identify problem participants who were not complying with the task. Fourth, as described previously, all participants completed an audio headphone check prior to completing the main experiment tasks, in which they were to indicate which of three tones was the quietest.

### 2.3 Stimuli and Experimental Tasks

**Auditory stimuli.** Taken from Choi et al., (2020), the speech stream was comprised of 12 artificially synthesized speech syllables, concatenated to create 4 trisyllabic "nonsense" words: *pau-to-ne*, *mai-pu-ki*, *nu-ra-fi*, *ga-mi-lu*. Syllables were generated using a female English speaker voice, with neutral intonation and were presented at a sampling rate of 44100 Hz, with no co-articulation. There were 93 syllables presented in each 30 second trial, and participants

were presented with 372 words total during the exposure period. The four trisyllabic nonsense words were continuously presented in random order (with the restriction that words, as well as two consecutive words, could not repeat back to back), allowing adjacent syllables within words to have a transitional probability of 1.0, and adjacent syllables between words to have a transitional probability of 0.33, leaving the statistical structure to be the prime word boundary cue.

**Visual distractor (multiple object tracking).** The classic multiple object tracking (MOT) paradigm as first developed by Pylyshyn and Storm (1988), was used to distract participants in the visual domain. During each trial, participants were to track a subset of dots (either 2 or 3 out of a total of 8) for 30 seconds, and indicate their positions via keypress at the end of the trial. Dots followed a Brownian motion. A sample of the task can be found in the appendix. During the 30 second trial, participants simultaneously heard the nonsense speech stream, and this speech stream stopped during the response period. Participants completed a total of 12 experimental trials, and 3 practice trials. During the practice trials, participants were to track 2 out of the 8 dots. If participants successfully tracked all 6 practice trial targets, the experiment began with a 3-target trial, and if participants missed one or more targets during the practice trials, the experiment began with a 2-target trial. To minimize the influence of varying baseline performance on the MOT task, a dynamic feedback system was implemented for the 12 main experimental trials. If participants obtained 100% accuracy on three consecutive 2-target trials, this was subsequently followed by a 3-target trial, and if participants obtained <100% accuracy on three consecutive 3-target trials, this was subsequently followed by a 2-target trial. Please see the appendix for sample still images of the MOT task.

Divided-attention participants were instructed to track the targets and ignore the sounds. Upon cessation of each trial, each dot was shown next to a corresponding number, and remained motionless in its current position. Participants were asked to type in the corresponding numbers of the targets. Full attention controls were instructed to pay attention to the sounds and ignore the moving dot visuals while still viewing the screen. During the response period, two random numbers appeared on the screen and participants were instructed to type them in, controlling for the motor response required of divided-attention participants.

**Linguistic distractor (self-paced reading).** Divided-attention participants completed a self-paced reading task in which sentences were presented on a screen one word at a time, with presentation of the next word controlled via keypress. Sentences were taken from the Harvard Sentence Inventory, a set of phonetically-balanced standardized sentences originally developed for speech quality research. Divided-attention participants were instructed to read sentences as fast as they could while paying attention to the content of the sentences, and to ignore the sounds. After each 30-second trial, participants answered comprehension questions corresponding to two randomly selected sentences that were read by the participant. Comprehension questions had a free response format (as opposed to multiple choice) to encourage effortful processing during reading. Sample sentence and comprehension question pairings include: Next Tuesday we must vote- *What day next week must we vote?* (answer: Tuesday); The zones merge in the central part of town- *What merges in the central part of town?* (answer: zones).

Full attention controls on the other hand were presented with randomly selected words from the Harvard Sentence Inventory, also presented one at a time, and were instructed to pay attention to the sounds. Words were presented in a random order to preserve the physical



integrity of the sentence stimuli allowing us to control for physical stimulus effects while also preventing the inclination to read the words for comprehension. Scrambled words were presented at a rate equivalent to the average word reading rate of the divided-attention participants (word rate presentation= 642ms). To control for the motor response required of divided-attention participants, full attention controls were asked to type in a random word displayed on the screen during the response period. Please see the Appendix for sample still images of the paradigm.

### 2.3.2 Statistical Learning Measures

**Familiarity Rating Task.** Following Batterink and Paller (2019), we assessed explicit statistical learning by administering a familiarity rating task in which participants were to provide an explicit familiarity judgment for single items presented on each trial. Items consisted of (1) trisyllabic words from the nonsense language, (2) part-word foils, consisting of a syllable from one word and two adjacent syllables from another word such as *MAI-to-ne*, or *ga-mi-FI*, and (3) non-word foils, consisting of syllables from three different words (and thus which never occurred together) such as *KI-PAU-nu*. Upon auditory presentation of the word/foil item, ratings were given via keypress on a 1-4 scale: 1- very unfamiliar, 2- unfamiliar, 3- familiar, 4- very familiar. Four part-words, 4 non-words, and the 4 nonsense language words were presented in random order, yielding a total of 12 trials. Response options were available for as long as the participant needed.

**Target detection task (TDT).** As an implicit measure of statistical learning, a speeded reaction time-based task was administered to capture knowledge that may have not been consciously accessible during the explicit recognition-based rating task. Participants were presented with an average of 14.45 ms long snippets of the nonsense speech stream, and were instructed to make a

keypress as quickly and accurately as possible in response to a specified target syllable. As found previously (e.g. Batterink & Paller, 2019; Batterink & Paller, 2017), learning is expected to be reflected by a graded reaction time effect whereby syllables occurring in later positions of the word would yield the fastest reaction times, as these syllables are more predictable than syllables occurring in word onset positions. That is, word-final syllables should yield the fastest reaction time, while word-initial syllables should yield the slowest reaction time, and word-medial syllables should yield intermediary reaction times.

Each of the 12 syllables from the nonsense language served as a target syllable 3 times, yielding a total of 36 streams. The target syllable occurred in each stream a total of 4 times, resulting in a total of 144 targets across the streams, with 48 targets per syllable position (1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>). Participants were first presented with a written form as well as sample audio of the target syllable. Once participants were familiarized with the target syllable, participants were to press a button to start the trial. The target syllable remained in written form on the screen throughout the duration of each trial. Responses not made within 1200ms of target onset were considered false alarms.

**Post-task interview.** Participants completed a post-task interview to further assess their awareness of the structure of the nonsense language, and to comment on the level of perceived effort on the distractor tasks. To assess awareness of the nonsense language, questions were first asked in a more general, open-ended format (ex. *did you notice anything interesting or unusual about the sounds?*) and gradually alluded to a hidden structure (ex. *did you notice any kind of pattern or structure in the sounds?*). Participants were also given a chance to elaborate on how difficult they found the tasks, as well as how often they found themselves mind-wandering.

## 2.4 Statistical Analyses

*Familiarity Rating task.* For each participant, ratings were computed for words, part-words, and non-words. Following previous findings (e.g. Batterink & Paller, 2017; 2019), we expected the highest familiarity ratings for words, intermediate ratings for part-words, and the lowest ratings for non-words. Ratings were analyzed with a repeated-measures ANOVA with word category (word, part-word, non-word) as a within-participants factor, and group (divided-attention visual, divided-attention reading, full-attention visual control, and full-attention reading control) as a between-subjects factor. Planned linear contrasts were used to compare the slope across word categories, between groups.

Though divided-attention participants may show significantly lower word ratings as compared to controls, they may nonetheless show significant evidence of word learning. To address this possibility, we conducted a repeated measures ANOVA within each group, as well as planned linear contrasts to compare the slopes between word ratings and non-word ratings. A significant linear contrast, reflecting a significant difference between word ratings and non-word ratings, would reflect learning. For correlational analyses, we took a composite rating task score for each participant by subtracting the average of the non-word and part-word rating scores from the average word rating score. Perfect sensitivity on this measure would be reflected by a score of 3, with values  $>0$  indicating learning.

*Target Detection Task.* Following a similar criterion in previous studies (e.g. Batterink & Paller, 2017; 2019), responses occurring within 1200ms were considered hits while those occurring outside the window of 0-1200ms post-target were considered misses. False alarms were calculated by subtracting the number of valid keypresses from the total number of keypresses for

that trial. RTs were expected to decrease as a function of syllable position, as later word syllables would be predicted by the preceding syllables. Average RTs for word-initial, word-medial, and word-final syllables were calculated for each participant and were subjected to a repeated-measures ANOVA with syllable position (initial, medial, and final) as a within-subjects factor, and group (divided-attention visual, divided-attention reading, full-attention visual control, full-attention reading control) as a between-subjects factor. Planned linear contrasts were used to examine group differences in slopes across syllable positions. As it is possible for divided-attention groups to show learning despite showing significantly reduced learning compared to controls, we conducted a repeated measures ANOVA within each group, with syllable position as a within-subjects factor, including planned linear contrasts to examine the slopes across syllable positions. A significant negative linear contrast would indicate whether reaction times significantly decreased as a function of syllable position, reflecting learning.

Since learning can occur during the TDT itself, we performed the above analysis again on the first third portion of the TDT data in order to confirm that the results found in the main analysis were replicated early on in the data. It is possible that group differences may be present immediately after exposure to the speech stream, but that these differences may be “washed out” due to the learning that can occur during the progression of the task since participants are re-exposed to the speech stream. Thus, we performed the same analysis again on the first third portion of the TDT data to confirm that the effects found on the entire analysis were already present early on in the task.

As a composite measure of learning as well as to account for baseline RT differences, a “RT prediction effect” was taken for each participant by dividing the difference in RT between word-initial and word-final syllables, by the word-initial reaction time  $((RT1-RT3)/RT1)$ . Greater

implicit word knowledge would be reflected by a larger RT prediction effect. This composite measure of learning on the TDT was used in our Pearson's R correlational analysis, in order to examine the relationship between the familiarity word rating and TDT performance.

## Chapter 3

### 3 Results

#### 3.1 Rating Task

Across all groups, words were rated as the most familiar ( $M=3.06$ ,  $SD=0.59$ ), followed by part-words ( $M=2.62$ ,  $SD=0.60$ ), and non-words ( $M=2.43$ ,  $SD=0.56$ ) as the least familiar, producing a significant effect of word type ( $F(2,220)=70.12$ ,  $p<.001$ ), linear contrast ( $t(220)=11.54$ ,  $p<.001$ ). Familiarity ratings differed significantly among the groups (Word type x Group:  $F(6,214)=2.74$ ,  $p=.014$ ). Planned linear contrasts were conducted to compare the word-partword-nonword rating performance slopes between: (1) divided-attention visual participants vs visual controls, (2) divided-attention reading participants vs reading controls, and (3) visual controls vs reading controls. These planned contrasts revealed that visual controls did not perform significantly differently from divided-attention visual participants ( $t(214)=1.41$ ,  $p=0.16$ ). In contrast, divided-attention reading participants showed significantly reduced learning performance as compared to reading controls ( $t(214)=3.13$ ,  $p=0.002$ ). In addition (and unexpectedly), reading controls showed significantly better performance than visual controls ( $t(214)=2.04$ ,  $p=0.040$ ).

In addition, we tested whether each of the four groups independently showed a significant evidence of learning, as indicated by a significant linear contrast for word type, conducted as part of a repeated-measures ANOVA. All groups showed a significant effect of word type, revealing a significantly higher rating for words as compared to non-words (*visual controls*: word type,

( $F(2,58)= 16.60, p<.001$ ); linear contrast, ( $t(58)= 5.41, p<.001$ ); *reading controls*: word type, ( $F(2,58)= 50.71, p<.001$ , linear contrast, ( $t(58)= 10.07, p<.001$ ); *divided-attention visual*: word type ( $F(2,48)= 27.21, p<.001$ ), linear contrast, ( $t(48)= 6.86, p<.001$ ); *divided-attention reading*: word type ( $F(2,50)= 4.16, p=0.02$ ), linear contrast, ( $t(50)= 2.74, p=0.009$ ).

In summary, on the Rating Task, participants in the divided-attention reading condition showed reduced performance relative to the control group, indicating a role of attention in SL. Nonetheless, significant SL was found within each of the four groups, suggesting that SL can proceed to some extent even outside the focus of directed attention.

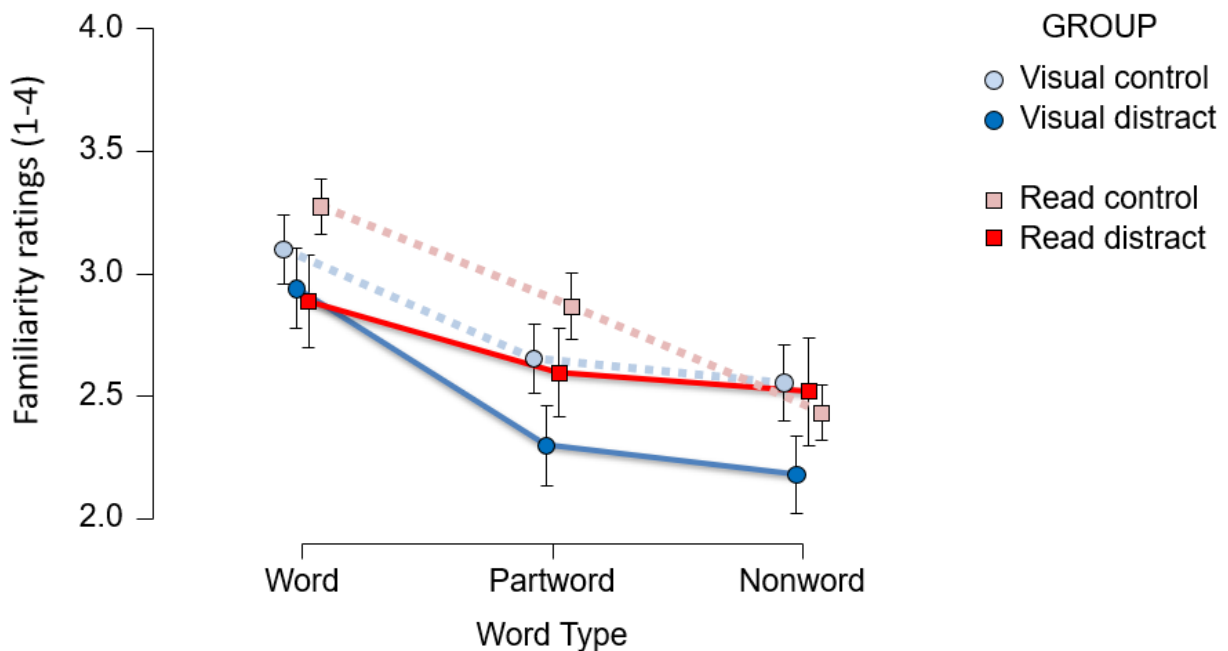


Figure 2. Mean rating scores for words, partwords, and nonwords on the explicit familiarity rating task across groups. Error bars denote the bootstrapped 95% CI.

### 3.2.1 Target Detection

Accuracy was high overall (hits:  $M= 89.13$ ,  $SD= 9.18$ ; false alarms:  $M= 12.87$ ,  $SD= 8.47$ ; misses:  $M= 15.60$ ,  $SD= 9.24$ ), with no significant differences in detection performance between groups (hits:  $F(3,107)= 0.93$ ,  $p= 0.43$ ; false alarms:  $F(3,107)= 1.04$ ,  $p= 0.38$ ; misses:  $F(3,107)= 0.94$ ,  $p= 0.43$ ).

Across all groups, RTs were the slowest for word-initial syllables ( $M= 559.29$ ms,  $SD= 85.69$ ms), intermediate for word-medial syllables ( $M= 536.36$ ,  $SD= 82.79$ ), and fastest for word-final syllables ( $M=503.42$ ,  $SD= 84.914$ ); Effect of Syllable Position: ( $F(2,220)= 94.61$ ,  $p<.001$ ), linear contrast: ( $t(220)=13.68$ ,  $p<.001$ ). This reduction in reaction time for more predictable syllables occurring later in the word reflects statistical learning across all groups.

RTs did not differ significantly across group, either overall across all positions (Group effect:  $F(3,107)= 0.60$ ,  $p=0.62$ ), or as a function of syllable position (Syllable Position X Group:  $F(6,214)= 1.45$ ,  $p=0.20$ ). Further, planned linear contrasts indicated no significant difference in RT slopes between visual controls and the divided-attention visual group ( $t(214)= -1.61$ ,  $p=0.11$ ), nor between the visual and reading controls ( $t(214)= -0.76$ ,  $p= 0.45$ ), nor between the divided-attention reading and reading controls ( $t(214)= -0.40$ ,  $p=0.69$ ), suggesting comparable learning among all four groups. To determine whether each group on their own showed significant evidence of learning, a repeated-measures ANOVA was conducted within each group, along with a linear contrast for syllable position. All groups showed a significant effect of syllable position as well as a significant linear contrast, indicating learning in all groups (*visual controls*: Syllable position effect: ( $F(2,58)= 33.36$ ,  $p<.001$ ), linear contrast: ( $t(58)= 7.87$ ,  $p<.001$ ); *divided-attention visual*: Syllable position effect: ( $F(2,48)= 17.52$ ,  $p<.001$ ), linear contrast: ( $t(48)= 5.82$ ,  $p<.001$ ); *reading controls*: Syllable position effect: ( $F(2,58)= 27.62$ ,

$p < .001$ ) linear contrast: ( $t(58) = 7.43, p < .001$ ); *divided-attention reading*: Syllable position effect: ( $F(2,50) = 20.92, p < .001$ ), custom contrast: ( $t(50) = 6.35, p < .001$ ).

Given the null group-related findings, a Bayesian repeated-measures ANOVA was conducted, revealing evidence in favour of a null effect of group ( $BF_{10} = 0.25$ ) and a null effect of group x position ( $BF_{10} = 0.15$ ).

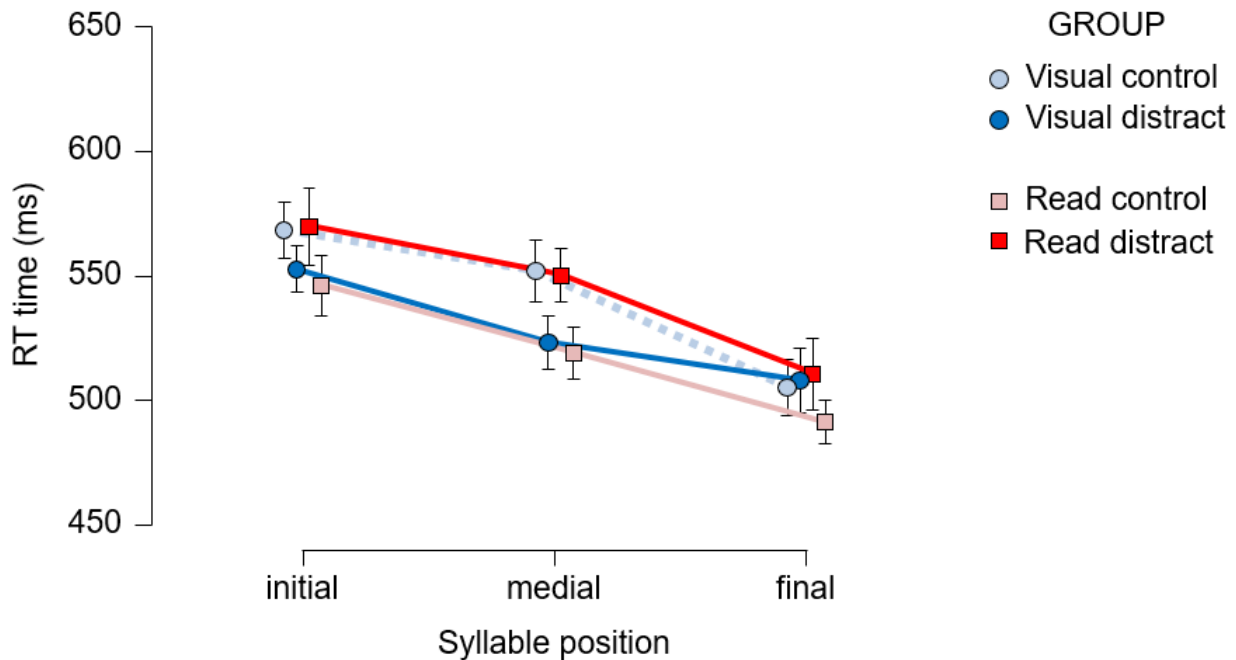


Figure 3. Mean reaction times (ms) for word-initial, word-medial, and word-final syllables across groups. Error bars denote the bootstrapped 95% CI.

### 3.2.2 Target Detection- First Third of Experimental Task

As learning can occur during the TDT task itself (due to additional exposure to the artificial language accrued during completion of the task), we reran the same analyses on only the first third trials (12 streams out of 36 total streams) in order to test whether the results found in the main TDT analysis would be replicated in the first portion. If we are unable to replicate the



same analysis, we would strongly suspect that a lack of group differences found using the entire TDT data may have been due to learning that occurs during the TDT task itself. On the other hand, if the same effects exist for both the first third portion and entire TDT task, we can be more confident that a lack of group differences were not due to effects being “washed out” with the learning that can occur during the re-exposure to the speech on the TDT.

Across all groups, RTs were greatest for word-initial syllables ( $M= 544.53$ ,  $SD= 90.26$ ), intermediate for word-medial syllables ( $M= 526.38$ ,  $SD= 96.80$ ), and fastest for word-final syllables ( $M= 486.52$ ,  $SD= 83.37$ ); effect of Syllable Position: ( $F(2,210)= 38.56$ ,  $p<.001$ ).

RTs did not significantly differ between group; Syllable Position X Group:  $F(2,210)= 1.05$ ,  $p= 0.40$ ). Planned contrasts indicated no significant difference in slopes between visual controls and the visual divided-attention group ( $t(210)= -1.04$ ,  $p=0.30$ ), nor between the two control groups ( $t(210)= -0.81$ ,  $p=0.42$ ), or between the reading controls and the reading divided-attention group ( $t(210)= -1.11$ ,  $p= 0.27$ ). These results replicate the main analysis, confirming no differences between the groups on this task.

### 3.2.3 Reaction Time Prediction Effect

Consistent with findings from the main RT analyses, the RT percent decrease effect did not significantly differ between groups  $F(3,107)= 0.77$ ,  $p=0.52$ ). In summary, participants in all four groups showed a comparable RT percent decrease, indicating that all groups showed similar SL performance on our implicit measure of SL, even when attention was not directed to the speech stream.

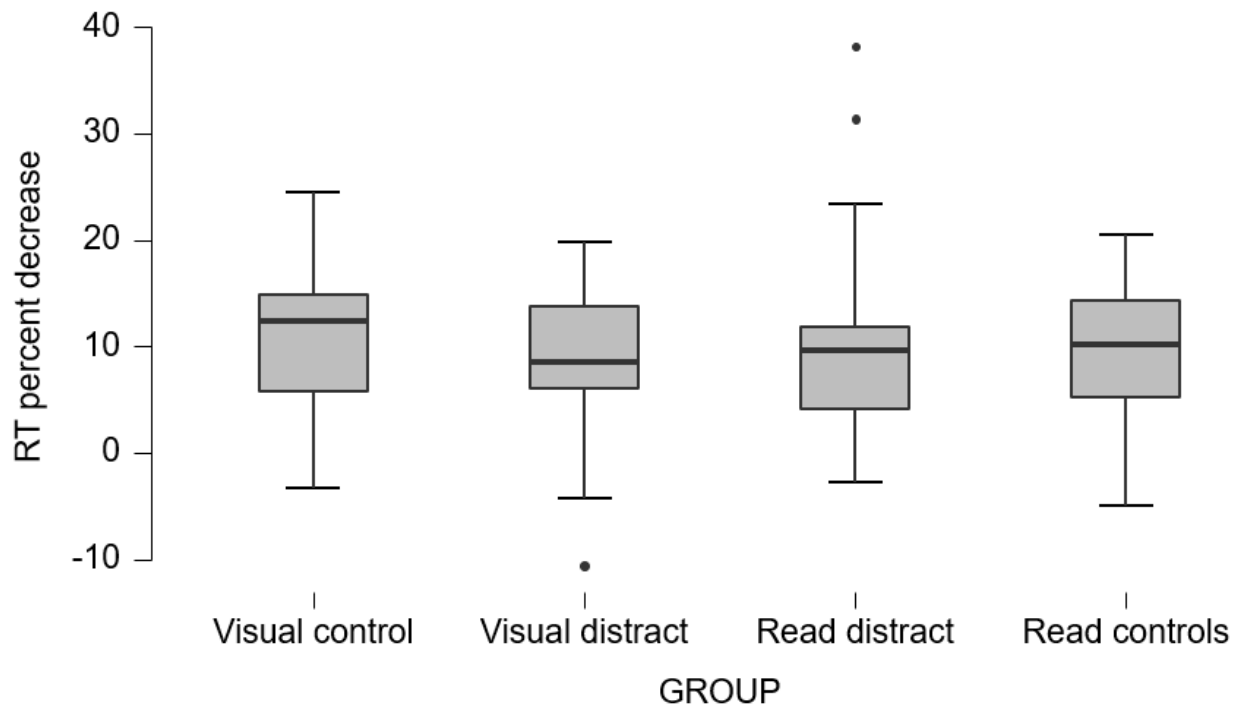


Figure 4. Boxplot showing the average RT percent decrease from first to third (last) syllables ( $RT_{\text{initial}} - RT_{\text{final}} / RT_{\text{initial}}$ ).

### 3.4 Accuracy on Concurrent Tasks (Divided Attention Groups only)

Participants had an overall accuracy of 81.70% ( $SD=0.41$ ) on the multiple object tracking task. Tracking performance is known to be fairly high on this task, with previous studies reporting tracking accuracy to be as high as 92.6% (Doran and Hoffman, 2010). Though tracking accuracy on our task was found to be lower than previous work, tracking accuracy has been shown to decline with an increasing number of targets (e.g. Drew et al., 2011) and distractors (e.g. Sears & Pylyshyn, 2000). On the self-paced reading task, participants answered an average of 65.70% ( $SD=0.12$ ) of comprehension questions correctly, which is an expected accuracy level given the difficulty of the task (Marsden et al., 2018).

### 3.5 Correlation Between Familiarity-Rating and Target Detection Task

Across all participants, the familiarity rating score for words and the RT prediction effect were not found to be significantly correlated ( $r=-.022, p=0.82$ ), consistent with some prior work (Batterink et al., 2015) and suggesting that these two measures reflect dissociable sources of knowledge accrued during SL.

### 3.6 Post-Task Questionnaire

A post-task questionnaire was administered to identify participants who had major technical problems, or major issues with complying to the task, as well as elaborate on the difficulty level of the attention manipulation tasks. Although a formal qualitative analysis was not conducted, anecdotally, the vast majority of participants who completed the post-task questionnaire commented that the MOT task required a lot of effort. The vast majority of participants who performed the reading task also commented that the task required a lot of effort. Interestingly, approximately 26 responders in the divided-attention reading group specifically alluded to the difficulty being related to the sounds in the background. One person wrote [the task took] *“a lot of effort because I needed to focus energy on drowning out the nonsense words from jumbling the worst I was reading up so that I could retain my reading comprehension.”* The vast majority of full-attention controls reported that they pay attention to the sounds.

## Chapter 4

### 4 Discussion

We sought to contribute to a literature where a consensus has not yet emerged regarding the role of top-down focused attention on statistical learning (SL) of speech. While work in the

area has largely relied on the use of an explicit recognition measure of learning (e.g. Lopez-Barroso et al., 2011; Palmer & Mattys, 2016; Toro et al., 2005), we have included an implicit reaction-time based measure of learning in addition to our explicit rating task measure. As well, our attention manipulations were designed to maximally tax two different types of limited-capacity resources (linguistic versus visual-spatial), allowing us to examine the effects of domain-specific and domain-general attentional resources on the SL of speech sounds. With our range of distractors and learning measures, our overarching aim was to investigate the impact of taxing top-down attention on SL. Our specific goals were to answer the following questions: (1) Does a reduction in focused attentional resources result in reduced SL? (2) Can SL occur even when attention is divided? (3) Is there a differential impact on learning based on the domain in which attention is taxed (language-based vs visual resources)?

On our explicit measure of learning, we found that taxing language-based resources resulted in poorer SL performance, although some degree of learning still occurred. In contrast, taxing visual resources did not significantly impact explicit SL performance on this measure. On our more implicit measure of learning, all groups followed a similar pattern of faster reaction times for later syllable positions, with no significant differences between groups, showing that implicit knowledge was not affected when focused attentional resources were taxed.

#### 4.1 Attentional Impact when Phonological-Linguistic Resources are Taxed

On our explicit rating measure, we found an attentional effect on learning specific to our divided-attention reading group, who showed significantly lower rating performance as compared to reading controls. In contrast, divided-attention visual participants performed similarly to visual controls, both showing above-chance learning. These results suggest that explicit SL is impaired only when language-based resources are taxed, while proceeding

uncompromised when visual resources are diverted. Additionally and unexpectedly, control groups performed significantly different from one another such that reading controls showed significantly better performance than visual controls. As both groups were instructed to pay attention to the speech sounds, the reason for this difference is not clear, although one possibility is that the presence of words on the screen placed reading controls into a more “verbal mode”, encouraging the processing of the nonsense speech stimuli.

Consistent with our hypothesis of a greater attentional effect for our language-based distractor as compared to our visual distractor, the divided-attention reading group showed reduced SL as compared to controls. In contrast, and similar to Batterink and Paller (2019), we found divided-attention visual participants and full attention participants to perform similarly on the rating task. These results fall in line with the classic finding of within-domain distractors bearing a greater attentional cost than across-domain distractors (e.g. Bayliss et al., 2003; Meiser & Klauer, 1999). A key cognitive process underlying both reading and SL, that was not present in our visual attention manipulation, is phonological processing, defined as “any task that requires awareness and/or the manipulation of the phonological structure of language” (Nittrouer, 1999). Reading involves discriminating between phonemes, and mapping out the phonological structure of a word to its orthographic representation. Similarly, statistically-driven speech segmentation involves phoneme discrimination, in order to make judgments about the phonological structure of words. The relationship between phonological processing and reading abilities has been illustrated in children with reading difficulties who have been shown to have difficulties identifying the different phonemes in a word, or adding and deleting phonemes from words (Lundberg, et al., 1980; Stanovich, 1986) For instance, phonological processing has been consistently implicated in those who have a learning disability specific to reading comprehension

and word recognition impairments (e.g. Olson et al., 1990; Snowling, 1981; Stanovich & Siegel, 1994). Previous research has also suggested that phonological processing plays a role in SL. For example, Lopez-Barroso (2011) and colleagues demonstrated that SL is significantly reduced when the articulatory rehearsal subcomponent of phonological working memory is taxed, by having participants continuously utter “blah” when hearing the nonsense speech stream. Similarly, Palmer & Mattys (2016) also reported impaired SL when participants performed a concurrent rhyming task, designed to tax phonological resources. As both the reading task and SL engage phonological processing, the reading task may have depleted this limited-capacity resource, resulting in impaired SL specific to the reading group.

Various theories concerning how phonological processing might contribute to SL have been proposed. Assaneo and colleagues (2019; 2020) show that those who show high spontaneous synchrony between a heard rhythm, and their own syllable production (termed “high synchronizers”) engage the fronto-parietal network when passively listening to a statistically structured nonsense speech stream, while those who do not display this spontaneous synchrony tendency (“low synchronizers”) do not engage this network. Further, high synchronizers showed greater SL word learning performance than low synchronizers, leading the authors to posit that the additional recruitment of the fronto-parietal network in high synchronizers contributes to SL.

Additionally, a previous study by Assaneo and colleagues (2019) further delineates the relationship between phonological processing, “high synchrony”, and SL, in their study demonstrating that high synchronizers show greater statistically-driven speech segmentation, and enhanced frontal brain synchrony with the isochronous auditory presentation of random syllables. These authors posit that enhanced synchrony in the frontal region with the heard

stimulus facilitates syllabic parsing due to the alignment of attention with their onset, which in turn can facilitate word learning. In high synchronizers, magnetoencephalography activity in this fronto-parietal network was found to align with the presentation of syllables, leading the authors to speculate that SL was boosted by enhancing syllable boundaries (Assaneo et al., 2019). In their later study, participants were instructed to interfere with the phonological loop by uttering a nonsense syllable during speech exposure, SL performance was reduced in only the high synchronizers— not low synchronizers— further suggesting a role of phonological processing in SL (2020). These findings corroborate research showing a link between auditory-motor synchrony and phonological processing by enhancing the detection of word boundaries (Woodruff Carr et al., 2014; see Tierney & Kraus, 2014 for a review).

Tierney & Kraus (2013) further speculate that high synchronization to a heard rhythm relies on the ability to discern fine-grained timing details in the speech stream, enabling the perceptual separation of different phonemes. Other researchers describe the contribution of phonological working memory to SL by highlighting its role in actively maintaining syllable composites (Palmer & Mattys, 2016). Palmer & Mattys (2016) show that SL was greater at a slower speech presentation rate, leading them to propose that the slower presentation rate allowed for more time for an “attentional refreshing” or rehearsal of syllable sequences, making them more likely to be stored in memory. However, given that auditory SL performance has been shown to be enhanced at faster presentation rates (Emberson et al., 2011), the nature by which phonological processing interacts with SL processes is not clear.

#### 4.2 Attentional impact on implicit vs explicit knowledge

Our language-based distractor reduced performance only on our explicit measure, without affecting performance on our implicit measure, consistent with the well-established

finding of implicit memory as more resilient as compared to explicit memory (e.g. Reber & Squire, 1994; Penfield and Milner, 1958; Tulving et al., 1982). Implicit memory is found to be resilient to factors such as brain damage (e.g. Graf et al., 1984, Cohen & Squire, 1980), divided-attention (e.g. Wolters & Prinsen, 1997), the passage of time, interference, (see Schachter et al., 1983 for a review). Our findings of an attentional effect for our reading group on the explicit measure and not the implicit measure as also in line the classic finding that explicit memory is impaired when attention is diverted at encoding ( Craik et al., 1996).

All four groups – including both divided-attention groups – showed significantly faster reaction times for later syllable positions, indicating significant facilitation to more predictable syllables as a result of SL. However, one limitation of the TDT is that it provides additional exposure to the language, such that additional SL may occur during the task itself. Thus, given that all four groups were asked to full attend to the speech stream segments during the TDT itself, it is possible that any effects of attention during the exposure period may have “washed out” with additional exposure to the language as the TDT progressed. However, we did not find evidence consistent with this possibility; our analysis restricted to the first third of the trials on this task replicated the main analysis, showing that the four groups did not significantly differ in performance even early on in the task, when RT effects resulting directly from learning during the exposure period should be the strongest. This finding suggests that the null effect of group is likely not due to the online learning that may occur during the TDT itself, although an additional control group is necessary to entirely rule out this explanation, a point that we return to below. Taken together, these findings suggest that implicit knowledge is not significantly impacted when attentional resources are taxed, while explicit knowledge is compromised only when



within-domain resources are taxed, supporting the idea that implicit memory may be more robust to attentional diversions (e.g. Prull, 2016).

In their review, Conway (2020) proposes that SL involves both implicit and explicit components, such that implicit forms of learning operate with minimal attentional requirements while more explicit aspects of learning require attentional resources. Supporting this idea, in a study by Batterink and colleagues (2015), participants showed learning on the implicit reaction time-measure, while a number of participants who showed learning on the implicit reaction time-based measure did not show learning on the explicit recognition task, suggesting that implicit learning may be more robust across participants as compared to explicit learning. They further propose that the implicit and explicit learning systems operate in parallel during SL tasks, such that the implicit learning system is always engaged, whereas the explicit memory system is optional, and perhaps only engaged in some participants and/or under certain learning conditions. Taken together, it is possible that while the implicit memory system was engaged for the divided-attention reading group, the reading task interfered with resources needed for explicit learning.

However, another possibility for our divergent findings on our explicit and implicit measures may simply be due to the nature of our implicit measure. As mentioned previously, the TDT re-exposes participants to the structured speech stream, providing an opportunity for additional learning. It is possible that disrupted learning for our divided-attention visual group was present immediately after the exposure period, but that these attentional effects were quickly washed away as learning continued to occur during the TDT itself. Future work will benefit from administering the TDT task to an additional group of control participants, who would complete this task without receiving any prior language exposure. This would provide a baseline measure

of performance attributable to learning on the task itself. If the four main experimental groups all show stronger implicit learning effects than this control group, this would strongly support the idea that implicit forms of SL can continue to operate in the absence of focused attention. In contrast, if the control group and four experimental groups perform equivalently, this would suggest that the equivalent learning effects we found for our four groups is entirely attributable to the learning that occurs on the TDT task itself.

### 4.3 Comparison to Past Findings

Our findings from the MOT demonstrated that SL proceeds unimpaired when visual resources are taxed, as assessed by both implicit and explicit measures. These results replicate those of Batterink & Paller (2019), who— using a slower paced, arguably less demanding visual attention manipulation— also found SL to occur on both their implicit and explicit measures. Our findings also follow that of Fernandes and colleagues (2010), who found that SL continued to occur for high TP words even when participants were taxed with a concurrent visual task. Taken together, these results continue to provide support for the relative automaticity of SL by showing learning can proceed even when attention is taxed with an extremely demanding visual task- as long as the task is visual, rather than phonological or language-based in nature.

Our findings from the reading task also converge with results by Lopez-Barroso and colleagues (2011) as well as Palmer & Mattys (2016), who found that taxing linguistic resources resulted in impaired, yet still above-chance learning on their explicit measure. Taken together, these findings supporting the notion that when within-domain resources are taxed, SL is impacted but can still occur to some extent.

In contrast, our findings are not fully in line with Toro et al. (2005), who found that SL was reduced to chance-level performance when auditory and visual attention resources were depleted. By comparison, we found above-chance learning on our explicit measure for both attentional manipulations. One possible reason for this discrepancy may be the differences in explicit measure, with Toro and colleagues (2005) using a 2-AFC task, whereas we used a familiarity rating task. Our task may have been more sensitive to relatively low levels of knowledge, as learners are able to provide a 1-4 rating, allowing for a finer-grained measure of word knowledge, and potentially capturing learning beyond what is normally captured by the 2-AFC task. In addition, one reason Toro and colleagues (2005) may have found SL when participants performed a concurrent auditory N-back task is that the concurrent auditory streams resulted in energetic masking, making the raw nonsense speech sounds perceptually unclear. Our reading task is performed silently, with a visual presentation of the reading stimuli, leaving the nonsense speech sounds perceptually clear and undegraded. One potential reason for the discrepancy in findings for our visual manipulations may be that Toro and colleagues used a more demanding visual manipulation, although we do not suspect this is the prime reason for the discrepancy in findings. They used nameable stimuli in their visual N-back task, making it possible for participants to have used a subvocal rehearsal strategy. This raises the possibility that linguistic resources were taxed, potentially accounting for the chance-level performance in their divided-attention visual group. As well, N-back performance is possible at the 3-back level (e.g. Batterink & Paller, 2019; Kirchner, 1958; Gevins et al., 1990), making it unlikely that attentional resources were maximally depleted in Toro and colleagues' study. At present, the discrepancy in results for the visual attention manipulations are still somewhat unclear. On the other hand, one similarity shared with our study is that Toro and colleagues (2005) found above-

chance SL when their visual N-back task was at a relatively slow presentation rate (750ms), providing support for the relative automaticity of SL.

Our results also do not perfectly align with those of those of Palmer & Mattys (2016) who found reduced SL on their explicit measure when visual resources were taxed, while we did not. One possibility is that the 2-AFC is more sensitive than the rating task in identifying group differences in performance. Another point to consider is their speech presentation rate manipulation- when speech presentation rate is taken into account, the discrepancy in results is not as drastic. These researchers presented nonsense speech stimuli either at a fast, normal, or slow rate and found that the attentional effect was strongest during a slow presentation rate. At a fast presentation rate, they report no attentional effect, and at a normal speech presentation rate – which is the most comparable to our speech presentation rate- they report only a marginal attentional effect. On the other hand, similar to our study, they also find above-chance learning when visual and linguistic resources are taxed, also showing that SL may occur to some extent when attention is taxed.

#### 4.4 Limitations and Future Directions

As mentioned previously, current work would benefit from the addition of a “no exposure” control group, who would complete the TDT with no prior language exposure. As learning can occur during the TDT itself, it is unclear what proportion of the learning effects observed in the current study are due to learning during the exposure phase versus during the TDT. Including this additional no-exposure control group will allow us to get an estimate of performance on this task with no prior learning at all.

Another major limitation includes the online mode of data collection. Participants completed the study remotely, and so there was no way to guarantee a uniformly, quiet, distraction-free, controlled environment across all participants. To control for this issue, we included multiple audio and attention checks throughout the study, such as our preliminary audio check, and the pause detection task for control participants. Nonetheless, it is possible that participants assigned to the full attention groups were not as fully engaged in listening to the audio streams as they would be in a controlled laboratory environment. If participants in the full attention conditions were somewhat distracted during the exposure period (for example, engaged in additional tasks or entertainment options), this would reduce the effects of our attention manipulations on our SL measures. As well, our study was subject to the usual drawbacks of online data collection such as possible participant fraud. Post-pandemic, future studies will take place in person where the researcher may monitor the participant, better control the study environment, and confirm participant engagement in the tasks.

Lastly, a discussion of an artificial language paradigm is incomplete without acknowledging the influence of prior language experience on learning outcomes (Siegelman et al., 2018). Though SL studies often assume a “blank slate” upon presentation of the artificial nonsense words, word composites are never truly “nonsense” to an individual with decades of language experience. Participants arrive with a lifetime of exposure to the statistical patterns and phonotactic properties of their native language, and this prior knowledge has been posited to impact SL performance, a concept known as “linguistic entrenchment” (Siegelman et al., 2018). It is possible that one or more of the trisyllabic words in our nonsense language may have inadvertently resembled English words, leading to idiosyncrasies in performance and/or greater learning as compared to words bearing less phonotactic resemblance. Future studies would

benefit by varying the word corpus for participants, either by randomly drawing unique sets of nonsense words from the syllable inventory, or using more than one artificial language.

To further understand and characterize the role of attention in SL, a major next step is to use a more sensitive EEG-based neural entrainment measure of learning, particularly in combination with the linguistic divided-attention manipulation, as this was the sole manipulation to result in reduced explicit knowledge. As described in the Introduction, traditional recognition measures of learning are sensitive to a number of secondary processes such as memory-related processes such as interference, retrieval, and decay, introspective abilities, and decision-making abilities. As well, they require on an overt behavioural response. Furthermore, recognition tasks and the TDT are post-learning measures, capturing the consequence of learning as opposed to the learning process itself. To overcome these limitations, our next study will include a more direct index of SL by capitalizing on the brain's tendency to oscillate at the frequency of a rhythmic external stimulus, known as *neural entrainment*. In a SL context, Batterink and Paller (2017) found that neural entrainment to the frequency of the embedded words in the speech stream increased as a function of exposure, and predicted performance on post-learning SL tests. In addition to bypassing the limitations of traditional post-learning measures, this direct index of SL also reveals the temporal dynamics of learning, which would allow us to investigate the process and time course of learning itself, rather than merely the downstream memory consequences of learning. With this neural entrainment measure, we plan to investigate the possibility that taxing linguistic attentional resources does not completely abolish SL, but rather slows down the progression of learning.

## 4.5 Conclusions

In the current study we report a disruption to SL learning in our divided-attention reading group. In contrast, SL continues to occur at normal levels even when participants are distracted with a demanding visual concurrent task. At a practical level, these results suggest that L2 learners may benefit from passively listening to the target language in the background without the need to consciously focus their attention and effort to the speech stream. We only have so much time in the day, and attention is a limited-capacity resource. Our finding that SL can still occur to some extent even in the presence of a highly demanding visual task may allow L2 learners to capitalize on periods of the day that would not normally be used for language learning.

However, since we found reduced SL when linguistic resources were taxed, L2 learners would be advised to avoid engaging in a secondary language-related visual tasks, such as reading, texting, and writing, when listening to the to-be-learned target language. On the other hand, learners may be able to participate in secondary visual or non-linguistic activities such as cleaning, knitting, putting together a puzzle, non-linguistic video games such as Tetris, and painting, while listening to the novel language and still successfully extract statistical regularities. A concurrent visual task may even be more helpful to learners over simply listening to the language with no concurrent task by curbing boredom and mind-wandering, possibly allowing them to engage in L2 listening for longer periods of time.

At a theoretical level, these results offer novel insights into the cognitive underpinnings of language acquisition by elucidating the attentional conditions in which SL can and cannot occur. We have contributed to the current literature with our study that suggests that SL can

indeed occur when a visual distractor is present, in contrast to a number of previous findings who have found a visual distractor to elicit a disruption to learning. Furthermore, these findings contribute to the growing notion of SL as involving multiple components (e.g. Arciuli, 2017; Emberson 2011; Siegelman et al., 2016), as we have illustrated differential effects for implicit and explicit aspects of SL, as well as different attentional effects when linguistic vs visual resources are taxed. Lastly, as SL is a domain-general phenomena extending to modalities such as vision and touch, these findings may inform questions of automaticity in areas beyond language.



## References

- Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track? Evidence for a resource-limited attentive tracking mechanisms. *Journal of Vision*, 7(13):14, 1–10, <http://journalofvision.org/7/13/14/>, doi: 10.1167/7.13.14.
- Assaneo, M. F., Orpella, J., Ripollés, P., Noeovich, L., López-Barroso, D., de Diego-Balaguer, R., & Poeppel, D. (2020). Population-level differences in the neural substrates supporting Statistical Learning. *bioRxiv*.
- Awh, E., Vogel, E. K., & Oh, S. H. (2006). Interactions between attention and working memory. *Neuroscience*, 139(1), 201-208.
- Baddeley, A.D. (1986). *Working memory*. Oxford University Press.
- Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105, 158–173.
- Baddeley, A. D., & Logie, R. H. (1999). *Working memory: The multiple-component model*.
- Baddeley, A. D., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language*, 27, 586–595.
- Baker, C. I., Olson, C. R., & Behrmann, M. (2004). Role of attention and perceptual grouping in visual statistical learning. *Psychological science*, 15(7), 460-466.
- 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. J., Reber, P. J., & Paller, K. A. (2015). Functional differences between statistical learning with and without explicit training. *Learning & Memory, 22*, 544-556.
- Batterink, L. J., & Paller, K. A. (2017). Online neural monitoring of statistical learning. *Cortex, 90*, 31-45.
- Batterink, L. J., & Paller, K. A. (2019). Statistical learning of speech regularities can occur outside the focus of attention. *Cortex, 115*, 56-71.
- Bayliss, D.M., Jarrod, C, Gunn, D.M., & Baddeley, A.D. (2003). The complexities of complex span: Explaining individual differences in working memory in children and adults. *Journal of Experimental Psychology: General, 132*, 71-92
- Bettencourt, K. C., & Somers, D. C. (2009). Effects of target enhancement and distractor suppression on multiple object tracking capacity. *Journal of vision, 9*(7), 9-9.
- Carr, K. W., White-Schwoch, T., Tierney, A. T., Strait, D. L., & Kraus, N. (2014). Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. *Proceedings of the National Academy of Sciences, 111*(40), 14559-14564.
- Cohen, M. A., Alvarez, G. A., & Nakayama, K. (2011). Natural-scene perception requires attention. *Psychological science, 22*(9), 1165–1172.  
<https://doi.org/10.1177/095679761141916>
- Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science, 210*(4466), 207-210.

- 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*(1), 24.
- Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychological science*, *17*(10), 905-912.
- de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*(1), 1-12. doi:10.3758/s13428-014-0458-y.
- De Diego-Balaguer, R., & Lopez-Barroso, D. (2010). Cognitive and neural mechanisms sustaining rule learning from speech. *Language Learning*, *60*, 151-187.
- Doran, M. M., & Hoffman, J. E. (2010). The role of visual attention in multiple object tracking: Evidence from ERPs. *Attention, Perception, & Psychophysics*, *72*(1), 33-52.
- 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.
- Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment.
- Fernandes, T., Kolinsky, R., & Ventura, P. (2010). The impact of attention load on the use of statistical information and coarticulation as speech segmentation cues. *Attention, Perception, & Psychophysics*, *72*(6), 1522-1532.

- Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3), 458-467. <http://dx.doi.org.proxy1.lib.uwo.ca/10.1037/0278-7393.28.3.458>
- Franco, A., Eberlen, J., Destrebecqz, A., Cleeremans, A., & Bertels, J. (2015). Rapid serial auditory presentation. *Experimental Psychology*, 62, 346–351.
- Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality vs. modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, 19, 117–125.
- Gevins, A. S., Bressler, S. L., Cutillo, B. A., Illes, J., Miller, J. C., Stern, J., et al. (1990). Effects of prolonged mental work on functional brain topography. *Electroencephalography and Clinical Neurophysiology*, 76(4), 339350.
- Graf, P., Shimamura, A., & Squire, L. R. (1985). Priming across modalities and priming across category levels: Extending the domain of preserved function in amnesia. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 11, 386-396.
- Graf, P., Squire, L. R., & Mandler, G. (1984). The information that amnesic patients do not forget. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(1), 164.
- Hendricks, M. A., Conway, C. M., & Kellogg, R. T. (2013). Using dual-task methodology to dissociate automatic from nonautomatic processes involved in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(5), 1491.

- Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*(3), 306-340.  
doi:<http://dx.doi.org.proxy1.lib.uwo.ca/10.1037/0096-3445.110.3.306>
- Janacek, K., & Nemeth, D. (2013). Implicit sequence learning and working memory: correlated or complicated?. *Cortex*, *49*(8), 2001-2006.
- Janacek, Karolina & Nemeth, Dezso. (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: A Journal Devoted to the Study of the Nervous System and Behavior*, *64*, 411-412. <https://doi.org/10.1016/j.cortex.2014.07.020>
- Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, *55*(4), 352-358.
- Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*(2), B35-B42.
- Kim, R., Seitz, A., Feenstra, H., & Shams, L. (2009). Testing assumptions of statistical learning: is it long-term and implicit?. *Neuroscience letters*, *461*(2), 145-149.
- Kubovy, M. (1988). Should we resist the seductiveness of the space:time::vision:audition analogy? *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 318–320
- Lopez-Barroso, D., de Diego-Balaguer, R., Cunillera, T., Camara, E., Münte, T. F., & Rodriguez-Fornells, A. (2011). Language learning under working memory constraints

- correlates with microstructural differences in the ventral language pathway. *Cerebral cortex*, 21(12), 2742-2750.
- Lundberg, I., Olofsson, A. & Wall, S. (1980). Reading and spelling skills in the first school years predicted phoneme awareness skills in kindergarten. *Scandinavian Journal of Psychology*, 21, 159–173.
- Marsden, E., Thompson, S., & Plonsky, L. (2018). A methodological synthesis of self-paced reading in second language research. *Applied Psycholinguistics*, 39(5), 861-904.  
doi:<http://dx.doi.org.proxy1.lib.uwo.ca/10.1017/S0142716418000036>
- Meyerhoff, H. S., Papenmeier, F., & Huff, M. (2017). Studying visual attention using the multiple object tracking paradigm: A tutorial review. *Attention, Perception, & Psychophysics*, 79(5), 1255-1274.
- Milner, B., Corkin, S., & Teuber, H. L. (1968). Further analysis of the hippocampal amnesic syndrome: 14-year follow-up study of HM. *Neuropsychologia*, 6(3), 215-234.
- Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive psychology*, 48(2), 127-162.
- Nittrouer S. (1999). Do temporal processing deficits cause phonological processing problems?. *Journal of speech, language, and hearing research : JSLHR*, 42(4), 925–942.  
<https://doi.org/10.1044/jslhr.4204.925>
- Olson, R. K., Wise, B., Conners, F., & Rack, J. P. (1990). Organization, heritability, and remediation of component word recognition and language skills in disabled readers. In T.

- H. Carr & B. A. Levy (Eds.), Reading and its development: Component skills approaches (pp. 261-322). San Diego, CA: Academic Press
- Palmer, S. D., & Mattys, S. L. (2016). Speech segmentation by statistical learning is supported by domain-general processes within working memory. *Quarterly Journal of Experimental Psychology*, *69*(12), 2390-2401.
- Palmer, S. D., Hutson, J., & Mattys, S. L. (2018). Statistical learning for speech segmentation: Age-related changes and underlying mechanisms. *Psychology and aging*, *33*(7), 1035.
- Penfield W, Milner B. Memory Deficit Produced by Bilateral Lesions in the Hippocampal Zone. *AMA Arch NeurPsych*. 1958;79(5):475–497.  
doi:10.1001/archneurpsyc.1958.02340050003001
- Prull, M. W., Lawless, C., Marshall, H. M., & Sherman, A. T. (2016). Effects of Divided Attention at Retrieval on Conceptual Implicit Memory. *Frontiers in psychology*, *7*, 5.  
<https://doi.org/10.3389/fpsyg.2016.00005>
- 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, P. J., & Squire, L. R. (1994). Parallel brain systems for learning with and without awareness. *Learning & memory*, *1*(4), 217-229.
- 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., 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. <https://doi.org/10.1111/j.1467-9280.1997.tb00690.x>
- 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., 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. [https://doi.org/10.1016/s0010-0277\(98\)00075-4](https://doi.org/10.1016/s0010-0277(98)00075-4)
- Schacter, D. L., Chiu, C. Y. P., & Ochsner, K. N. (1993). Implicit memory: A selective review. *Annual review of neuroscience*, 16(1), 159-182.
- Sears, C. R., & Pylyshyn, Z. W. (2000). Multiple object tracking and attentional processing. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 54(1), 1.
- 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 B: Biological Sciences*, 372(1711), 20160059. <https://doi.org/10.3758/s13428-016-0719-z>
- 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.



- Snowling, M. J. (1981). Phonemic deficits in developmental dyslexia. *Psychological Research*, 43, 219-234.
- 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.
- Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression based test of the phonological-core difference model. *Journal of Educational Psychology*, 86, 24-53.
- Stanovich, K. (1986). Cognitive processes and reading problems of learning disabled children: Evaluating the issue of specificity. In J. Torgesen & B. Wong (Eds.), *Phonological and educational perspectives in learning disabilities* (pp. 87–131). New York: Academic Press.
- Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, 7(1), 53-71
- Thiessen, E. D., & Saffran, J. R. (2007). Learning to learn: Infants' acquisition of stressbased strategies for word segmentation. *Language Learning and Development*, 3(1), 73-100.
- Toro, J. M., Sinnott, S., & Soto-Faraco, S. (2011). Generalizing linguistic structures under high attention demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(2), 493.
- Toro, J. M., Sinnott, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, 97(2), B25-B34.

- Treisman, A., & Davies, A. (2012). Divided attention to ear and eye. In J. Wolfe & L. Robertson (Eds.), *From perception to consciousness: Searching with Anne Treisman* (pp. 24–31). Oxford University Press. <https://doi.org/10.1093/acprof:osobl/9780199734337.003.0005>
- Tulving, E., Schacter, D. L., & Stark, H. A. (1982). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8(4), 336–342. <https://doi.org/10.1037/0278-7393.8.4.336>
- Unsworth, N. (2019). Individual differences in long-term memory. *Psychological Bulletin*, 145(1), 79.
- Wolters, G., & Prinsen, A. (1997). Full versus divided attention and implicit memory performance. *Memory & Cognition*, 25(6), 764-771.
- Woods, A. T., Velasco, C., Levitan, C. A., Wan, X., & Spence, C. (2015). Conducting perception research over the internet: a tutorial review. *PeerJ*, 3, e1058. <https://doi.org/10.7717/peerj.1058>
- Zhao J, Al-Aidroos N, Turk-Browne NB (2013) Attention is spontaneously biased toward regularities. *Psychol Sci* 24:667–677

## Appendices

### Appendix A: Ethics Approval



**Date:** 31 August 2020

**To:** Dr Laura Batterink

**Project ID:** 112359

**Study Title:** Behavioral and EEG studies of language learning

**Application Type:** NMREB Amendment Form

**Review Type:** Delegated

**Full Board Reporting Date:** 02/Oct/2020

**Date Approval Issued:** 31/Aug/2020 13:20

**REB Approval Expiry Date:** 24/Aug/2021

Dear Dr Laura Batterink,

The Western University Non-Medical Research Ethics Board (NMREB) has reviewed and approved the WREM application form for the amendment, as of the date noted above.

**Documents Approved:**

Document Name	Document Type	Document Date	Document Version
General Website Recruitment	Recruiting Advertisements	01/Aug/2020	2
MTurk Prolific recruitment	Recruitment Materials	01/Aug/2020	3
Research Plan Reflecting New Changes	Protocol	01/Aug/2020	2
Visual_stimuli	Other Materials	01/Aug/2020	1
Auditory_stimuli	Other Materials	01/Aug/2020	1
Familiarity_Rating	Other Materials	01/Aug/2020	1
reaction_time	Other Materials	01/Aug/2020	1
Online consent form	Implied Consent/Assent	18/Aug/2020	4

REB members involved in the research project do not participate in the review, discussion or decision.

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.

Please do not hesitate to contact us if you have any questions.

Sincerely,

Katelyn Harris, Research Ethics Officer on behalf of Dr. Randal Graham, NMREB Chair

**Note:** This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).

## Appendix B: Consent Form

**Project Title:** Behavioral and EEG Studies of Language Learning

**Principal Investigator:**

Dr. Laura Batterink

Department of Psychology, The University of Western Ontario, London, ON

**1. Invitation to Participate**

You are being invited to participate in a research study about how adults learn new languages.

The purpose of this letter is to provide you with information required for you to make an informed decision regarding participation in this research.

**2. Why is this study being done?**

The purpose of the study is to investigate how people pick up on different aspects of language, such as vocabulary and grammar. Compared to children, adults often struggle to acquire a new language. By understanding the neural mechanisms that adults rely on when they are exposed to a new language, we may better understand why adults often have more difficulty acquiring new languages. We may also get useful information about the best learning practices and types of training to help adults acquire a new language.

**3. How long will you be in this study?**

It is expected that this study will take approximately 45 minutes to 1.5 hours to

complete. **4. What are the study procedures?**

The experiments conducted as part of this study will test how humans process and learn about different types of linguistic stimuli, such as syllables, words, phrases and sentences. If you agree to participate, you will be asked to listen to language-related auditory stimuli and/or read words and sentences on a screen. You may be asked to perform different tasks associated with the stimuli, such as responding to targets by pressing a button, or making different judgments or ratings about your impressions of the stimuli. These tasks will be administered through the online data collection platform Pavlovia, which is a launch platform for online experiments, used widely in amongst behavioural science researchers.

**5. What are the risks and harms of participating in this study?**

Page 1 of 4 Version Date: August 18, 2020

There are no known or anticipated risks or discomforts associated with participating in this study. **6. What are the benefits?**

You do not directly stand to benefit from this study. Although you may not directly benefit from your participation, the information gathered may provide benefits to society as a whole which include enhancing our scientific understanding of language, learning, and the brain, and leading to advancements in second language training and treatment of language-related disorders (for example, specific language impairment and autism).

### **7. Can participants choose to leave the study?**

Participation in this study is voluntary. You may refuse to participate, refuse to answer any questions or withdraw from the study at any time with no effect on your future eligibility to participate in a research study.

### **8. How will participants' information be kept confidential?**

Information obtained from this study will be kept anonymous with no link to your personal data, although if you later decide to have your data withdrawn, this can be done if you provide us with your worker ID code. On the study platform, you will be assigned a unique, random participant ID. Identifiable information that will be collected during the study will include your partial date of birth (birth month and year), in order for us to look at age-related trends in the data. In the event of publication, any data resulting from your participation will be identified only by case number, without any reference to your personal information. While we do our best to protect your information, the inclusion of your partial date of birth may allow someone to identify you, although this is highly unlikely. Your data will be stored securely on servers administered by online experimental platforms, Qualtrics and Pavlovia.

Your survey responses will be collected through a secure online survey platform called Qualtrics. Qualtrics uses encryption technology and restricted access authorizations to protect all data collected. In Addition, Western's Qualtrics server is in Ireland, where privacy standards are maintained under the European Union safe harbour framework. The data will then be exported from Qualtrics and securely stored on Western University's server.

Pavlovia is launching platform for online experiments. Pavlovia is compliant with the General Data Protection Regulation (GDPR) principles, which means your data are anonymised as far as possible, as early as possible. If files are shared with other researchers or the results are made public, any personal information that could identify you will be removed. Only anonymized data will be shared outside the research team (e.g., in an open access repository for publication purposes, or for other researchers to verify the findings or re-analyze). Study records will be maintained for a minimum of 7 years and then will be securely deleted electronically. A list linking your study number with your participant ID will be kept by the researcher in a secure place. Your data may be retained indefinitely and could be used for future research purposes (e.g. to answer a new research question). By consenting to

participate in this study, you are agreeing that your data can be used beyond the purposes of this present study by either the current or other researchers.

Page 2 of 4 Version Date: August 18, 2020

Representatives of the University of Western Ontario Non-Medical Research Ethics Board may require access to your study-related records to monitor the conduct of the research.

## **9. Are participants compensated to be in this study?**

If you were recruited through SONA, you will receive 1.0 credit upon completion of this study.

If recruited through Prolific, you will receive the GBP equivalent of \$10 CDN per hour to participate in this study. If you do not complete the entire study you will still be compensated a pro-rated amount (based on the same rates specified above: \$10 CDN/h). When calculating prorated compensation, your total participation time will be rounded up to the nearest half hour. For example, if you withdraw after 1 hour and 15 minutes, your participation time will be rounded to 1.5 h and you will receive \$15CDN. Therefore, even if you withdraw prior to completing study, you will still be compensated for the amount of time you spent participating.

## **10. What are the rights of participants?**

Your participation in this study is voluntary. You may decide not to be in this study. Even if you consent to participate you have the right to not answer individual questions or to withdraw from the study at any time. If you are a student at Western and you choose not to participate or to leave the study at any time, it will have no effect on your academic standing.

You do not waive any legal right by signing this consent form

## **11. Whom do participants contact for questions?**

If you have questions about this research study please contact Laura Batterink, Principal Investigator

If you have any questions about your rights as a research participant or the conduct of this study, you may contact The Office of Human Research Ethics

This office oversees the ethical conduct of research studies and is not part of the study team. Note that everything you discuss will be kept confidential.

## **12. Consent**

If you would like to take part in the experiment, please complete the consent form, which can be found on the following webpage. Because you are completing this study online, you will be providing implied (not written) consent, which entails agreeing to a consent statement

through checking a box on the computer screen. By completing this consent form, you do not waive your legal rights nor release the investigator(s) and sponsors from their legal and professional responsibilities.

If you decide that you would not like to take part, simply exit out of your browser window.

Page 3 of 4 Version Date: August 18, 2020

**This letter is yours to keep for future reference.**

## Appendix C: Post-Task Questionnaires

Divided attention group (visual):

[https://uwo.eu.qualtrics.com/jfe/form/SV\\_1YQa80HpIFCG3uR](https://uwo.eu.qualtrics.com/jfe/form/SV_1YQa80HpIFCG3uR)

Questions: 1) You tracked the motion of balls while you listened to nonsense sounds. Did this task require a little or a lot of your effort or attention? Please describe.

2) Did you ever lose focus on tracking the balls and pay attention to the nonsense sounds? If so, approximately what proportion of the time were you paying attention to the sounds?

3) Did you notice anything interesting or unusual about the sounds? Please describe what you heard.

4) Did you notice any kind of pattern or structure in the sounds? If yes, please elaborate.

5) Did you have any technical issues? Please explain.

Divided attention group (reading)

[https://uwo.eu.qualtrics.com/jfe/form/SV\\_a9pioqM9ahN8k3X](https://uwo.eu.qualtrics.com/jfe/form/SV_a9pioqM9ahN8k3X)

Questions: 1) You read a series of sentences while you listened to nonsense sounds. Did this task require a little or a lot of your effort or attention? Please describe.

2) Did you ever lose focus on reading the sentences and pay attention to the nonsense sounds? If so, approximately what proportion of the time were you paying attention to the sounds?

3) Did you notice anything interesting or unusual about the sounds? Please describe what you heard.

4) Did you notice any kind of pattern or structure in the sounds? If yes, please elaborate.



5) Did you have any technical issues? Please explain.

Full attention control groups

[https://uwo.eu.qualtrics.com/jfe/form/SV\\_41v9bJwD4njfVUF](https://uwo.eu.qualtrics.com/jfe/form/SV_41v9bJwD4njfVUF)

Questions: 1) Did you pay attention to the sounds? If not, please explain why.

2) Did you ever lose focus on the sounds? If so, approximately what proportion of the time did you lose focus?

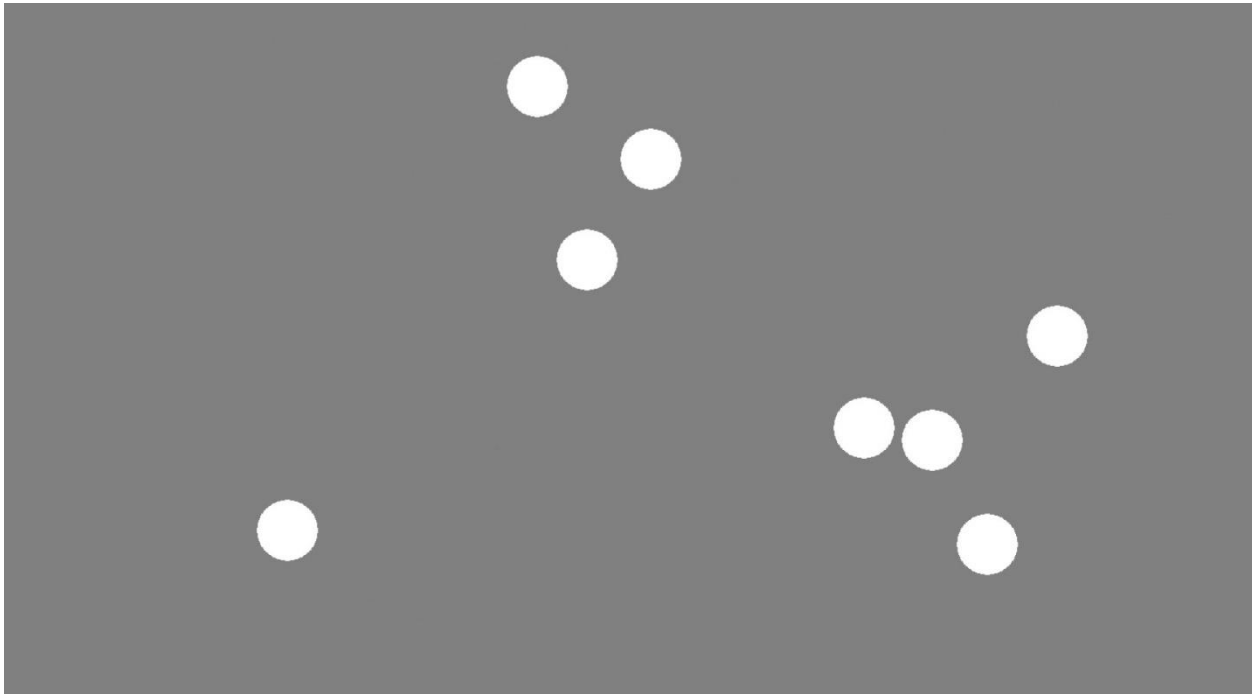
3) On a scale of 1-10, how interesting or engaging did you find the task? 1= not at all 10= very interesting or engaging

4) Did you notice anything interesting or unusual about the sounds? Please describe what you heard.

5) Did you notice any kind of pattern or structure in the sounds? If yes, please elaborate.

6) Did you have any technical issues? Please explain.

Appendix D: Multiple Object Tracking samples



Participants were to track 2-3 of the randomly moving balls

*Sample video of the MOT task:*

[https://www.youtube.com/watch?v=HIOVHF6mGjo&ab\\_channel=StaceyReyes](https://www.youtube.com/watch?v=HIOVHF6mGjo&ab_channel=StaceyReyes)

## Appendix E: Curriculum Vitae

Name: Stacey Reyes

**Post-secondary Education and Degrees:** Western University,  
London, Ontario, Canada  
2019-2021 M.Sc. Psychology Candidate

Queen's University  
Kingston, Ontario, Canada  
2015-2017 M.Sc. Psychology Candidate

York University  
Toronto, Ontario, Canada  
2010-2014, B.A. Psychology

University of Waterloo  
Waterloo, Ontario, Canada  
2009-2010, B.Sc. Candidate

**Honours and Awards:** Social Sciences and Humanities Research Council  
of Canada (SSHRC) Masters  
2015-2016

Tri-Agency Recipient Recognition Award  
2015

**Related Work Experience:** Research Assistant  
Cognitive Neuroscience & Sensorimotor  
Integration Laboratory  
University of Toronto Scarborough  
2019

Behavioural Coder  
Anxiety Research Clinic  
York University  
2013-2015

Research Assistant  
Children's Learning Projects Laboratory  
York University  
2012-2014

**Conference Presentations**

Reyes, S.D., Batterink, L.J. The influence of top-down attention on statistical learning of speech using an EEG measure of neural entrainment. *Cognitive Neuroscience Society*. March 2021. Virtual conference.

Reyes, S.D., Batterink, L.J. The influence of top-down attention on statistical learning of speech using an EEG measure of neural entrainment. *Society for the Neurobiology of Language*. October 2020. Virtual conference.

Reyes, S.D., M. Segers, & J.M. Bebko. The Role of Sensory Processing in Audiovisual Integration of Speech in Autism Spectrum Disorder. Poster Presented at the 44th Annual *Ontario Undergraduate Psychology Thesis Conference*. May 2014. Kingston, Ontario.

Reyes, S.D., M. Segers, & J.M. Bebko. The Role of Sensory Processing in Audiovisual Integration of Speech in Autism Spectrum Disorder. Poster Presented at the *York University High Scholars Event*. April 2014. Toronto, Ontario.

Reyes, S.D., M. Segers, & J.M. Bebko. The Role of Sensory Processing in Audiovisual Integration of Speech in Autism Spectrum Disorder. Poster Presented at the *Ontario Association on Developmental Disabilities Research Special Interest Group conference*. April 2014. Kingston, Ontario.

Reyes, S.D., M. Segers, & J.M. Bebko. The Role of Sensory Processing in Audiovisual Integration of Speech in Autism Spectrum Disorder. Poster Presented at the *York University Psychology Honours Thesis Conference*. April 2014. Toronto, Ontario.