Electrophysiological correlates of phonological and temporal regularities in speech processing

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Download date: 06 May. 2024

Electrophysiological correlates of phonological and temporal regularities in speech processing

Alexandra Katherine Emmendorfer

Electrophysiological correlates of phonological and temporal regularities in speech processing

Dissertation

To obtain the degree of Doctor at the Maastricht University,
on the authority of the Rector Magnificus Prof. dr. Pamela Habibović
in accordance with the decision of the Board of Deans,
to be defended in public
on Tuesday, 22nd of February, 2022 at 13:00 hours

by
Alexandra Katherine Emmendorfer

Supervisors
Prof. dr. Bernadette M. Jansma
Prof. dr. Sonja A. Kotz
Prof. dr. Milene Bonte

Assessment Committee
Prof. dr. Alexander Sack (Chair)
Prof. dr. Valéria Csépe (Hungarian Academy of Sciences)
Dr. Marie Lallier (Basque Center on Cognition, Brain and Language)
Dr. Michael Schwartze

Table of Contents

Chapter 1

General Introduction

Current theories of neural processing examine the brain as an organ that not only passively reacts to environmental events, but is also actively involved in generating predictions about upcoming input (e.g., Engel et al., 2001; Raichle, 2010; Rao & Ballard, 1999). These theories emerged from observations that our perception, as well as the brain's representation of sensory input, does not always correspond directly to the physical properties of the stimulus. Instead, the current sensory input is integrated with contextual cues, as well as prior knowledge of our environment to guide perception (Summerfield & de Lange, 2014). Such predictive capabilities of the brain can allow us to extrapolate from sensory data and internal states to anticipate upcoming input and adapt our behavior. For example, we can estimate the trajectory of an object that is moving toward us, and move our hand to catch it, or duck out of the way to avoid being struck. In a noisy environment, we can also make use of such predictions to better understand our conversation partner. While walking through the forest, we may anticipate the rustling of leaves beneath our feet but become startled when we hear rustling in the bushes behind us. Such predictive capacities are thought to contribute to a range of aspects of cognitive functioning, including motor control, perception and attention, as well as language. Disruption of the mechanisms underlying predictive processing may be associated with atypical skill development, resulting in impairments such as developmental dyslexia.

This chapter introduces the concept of predictive processes in the brain, with a focus on auditory and speech processing. First, a global overview of theoretical frameworks describing different types of predictions is provided, where we distinguish between predictions about the content and timing of sensory input, as well as predictions of externally generated and self-generated sensations. This is followed by examples specific to the domain of language, in speech perception and production, and how they vary in dyslexic readers. Finally, the electrophysiological correlates investigated in the context of predictive processing are presented.

1 Theoretical frameworks of predictive processing

A central task and challenge of perception is to generate percepts that are both accurate and informative (Press et al., 2020). A prominent theoretical framework proposed to solve this task is predictive coding. The central idea of this framework is that, rather than representing the sensory input directly, our brain instead represents

the deviation between expected and actual sensory input, or the prediction error (Aitchison & Lengyel, 2017; Friston, 2005; Rao & Ballard, 1999; Spratling, 2017). Such a mechanism may improve the efficiency of neural processing, as our brain no longer needs to allocate further resources for processing anticipated events. This implies that unexpected or surprising events should lead to greater neural activity and perceived intensity. To optimize efficiency in neural processing and behavior, predictions compared to the sensory input must be as accurate as possible, thereby minimizing the prediction error and the information that must be represented and transmitted for higher-level processing. To achieve this, our predictions must be informed by regularities in our environment and continuously updated when exposed to new information. A possible implementation of this is through hierarchical generative models and empirical Bayesian inference (Aitchison & Lengyel, 2017; Friston, 2005, 2009). This framework suggests that the brain attempts to infer the underlying cause of sensory input to minimize surprise (approximately equivalent to prediction error). It does so by optimizing the prediction, i.e., updating the internal representation of the world, or by optimizing behavior, to selectively sample more expected input.

In addition to anticipating the nature of external events, our brain also formulates predictions about the sensory consequences of our own actions (hereafter *motor prediction*). This procedure is described by the forward model of motor control (Wolpert & Miall, 1996). According to this framework, the motor cortex sends a copy of the motor command, termed efference copy, to the sensory cortices, representing the anticipated sensory outcome of an action. Like a prediction of external events, this anticipated outcome is compared to the actual sensory input, leading to a reduced sensory response commonly referred to as motor-induced suppression. A widely experienced example of this is the inability to tickle oneself (Blakemore et al., 2000). This mechanism facilitates the distinction between self- and externally generated events, as in the previously mentioned example of rustling leaves beneath our feet, vs. rustling leaves in the bushes behind us. It further allows us to monitor and correct our motor output more readily.

While these theories provide a useful representation of how we predict the content or quality of sensory input, they often neglect the aspect of temporal information. In addition to anticipating the formal structure of an upcoming event

(hereafter *formal prediction*), we also formulate predictions about their timing (hereafter *temporal prediction*). While formal and temporal predictions may interactively modulate perception and behavior (e.g., Schwartze et al., 2011), it is not necessary to generate a formal prediction to accurately anticipate the timing of an event, and vice versa. For example, when standing at a red light, we may accurately anticipate that it will turn green, even if we cannot say with certainty when this will happen. However, if this traffic light displays a countdown, we can be prepared to move as soon as the light turns green. Natural stimuli are often characterized by temporal regularities which allow us to efficiently direct our attention towards the next event to guide perception (Nobre & van Ede, 2018). Periodicity in the timing of events is typically associated with shorter reaction times and improved response sensitivity or accuracy (Ellis & Jones, 2010; Heynckes et al., 2020; Morillon et al., 2016), suggesting that this form of prediction is important for aligning our actions to external events, and highlighting a close link between the motor system and temporal processing (Kotz & Schwartze, 2011). However, a sequence of events does not need to be periodic to evoke temporal expectations. A predictable, but aperiodic, sequence of events may still facilitate perception (Heynckes et al., 2020; Lawrance et al., 2014; ten Oever et al., 2014). In fact, natural signals rarely are precisely periodic in nature. Such predictions may be associated with learned temporal contingencies between events, and would thus rely more on memory-based interval-timing rather than classically proposed entrainment models of temporal prediction (Rimmele et al., 2018).

2   Predictions in speech and language processing

Speech is a highly complex and variable signal, yet humans can extract meaning from this signal in a seemingly effortless manner. Predictive processing has been proposed as a mechanism that facilitates these dynamic computational demands in daily listening situations where the sensory input is suboptimal due to e.g., background noise, multiple speakers, or variations in accent. When listening to a speaker, we apply our prior knowledge as well as contextual cues including visual information such as gestures (ter Bekke et al., 2020) and movement of the articulators (Peelle & Sommers, 2015), and the rules of the language (DeLong et al., 2005; but see Nieuwland et al., 2018) to anticipate the upcoming content of an utterance. Predictions might be made

about upcoming words in a sentence, based on semantic context (Freunberger & Roehm, 2016; Kutas & Hillyard, 1980), to facilitate perception in noisy environments, and can even be exploited to facilitate language switching in multilingual settings (Martin et al., 2016). At the sublexical level (i.e., independent of lexical or semantic content), our brain is tuned toward regularities in combinations of speech sounds (Bonte et al., 2005; Di Liberto et al., 2019; Vidal et al., 2019) and speech rhythm (Kotz et al., 2018). In any given language, certain speech sounds are more likely to co-occur than others. The distributional probability of these combinations of speech sounds is termed phonotactic probability. Similarly, any given language will be characterized by regularities in fluctuations of the amplitude envelope, giving rise to the speech rhythm. One aspect of speech rhythm that can be characteristic of a language is the patterns of alternating stressed and unstressed syllables (Dauer, 1983; Ding et al., 2017; Pike, 1962). In this dissertation, the focus will lie on predictions of phonotactic probability and syllable stress as examples of formal and temporal regularities (Figure 1).
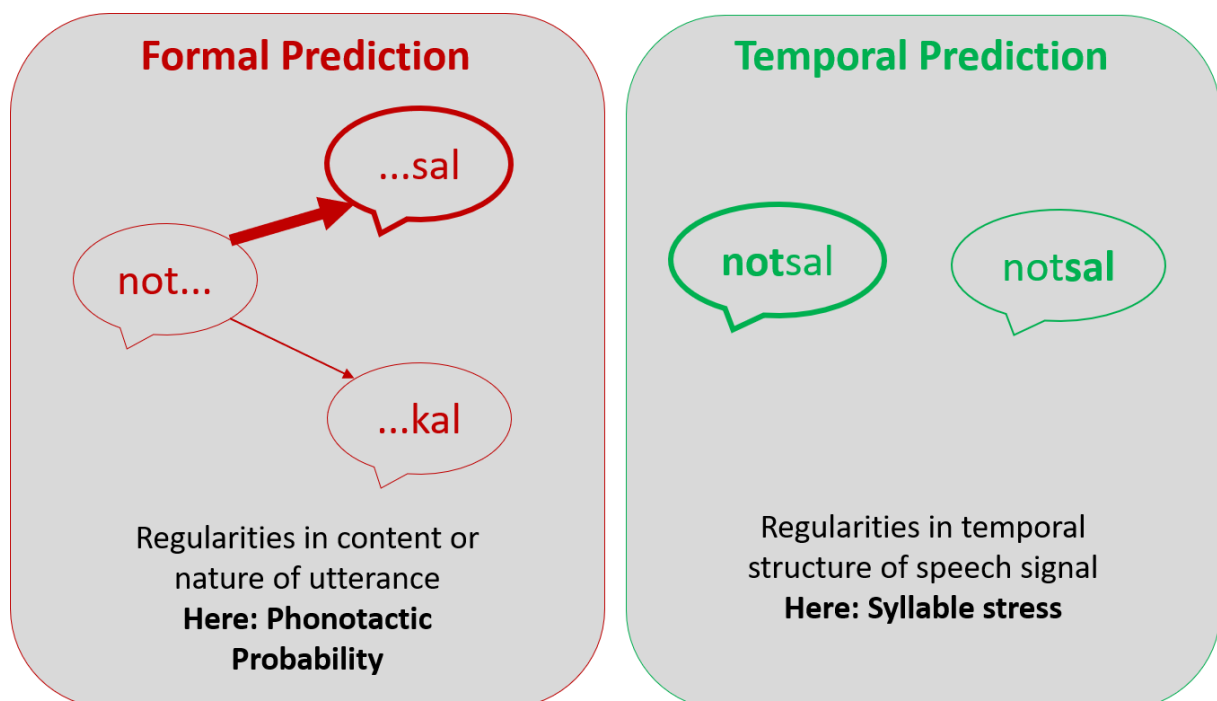
Figure 1: Phonotactic probability and syllable stress as examples of formal and temporal predictions in speech processing. Left: Pseudoword *notsal* has high phonotactic probability, while *notkal* has low phonotactic probability. Right: Bold font indicates stressed syllable, where first syllable stress is more common than second syllable stress in Dutch.

During development, we acquire implicit knowledge of these regularities, which has been demonstrated as early as 8 months of age for syllable co-occurrence (Saffran et al., 1996), and around 6-9 months of age for syllable stress patterns (Jusczyk et al., 1999). Distributional regularities in the speech signal are used during development to guide word segmentation/word learning and continue to facilitate sublexical language processes throughout the lifespan. Children and adolescents show an effect of phonotactic probability on nonword repetition (Edwards et al., 2004; Munson, Edwards, et al., 2005; Munson, Kurtz, et al., 2005), starting as early as 2 years of age (Verhagen et al., 2017), where high probability items are generally associated with improvements in both speed and accuracy. This effect may also be further mediated by vocabulary size: children with smaller vocabularies struggle more with low probability items, leading to a larger effect of phonotactic probability than observed in children with larger vocabularies (Edwards et al., 2004; Munson, Edwards, et al., 2005; Munson, Kurtz, et al., 2005). Interestingly, when it comes to word learning, the effect of phonotactic probability seems to be reversed. By analyzing a naturalistic database of words known by infants, Storkel (2009) revealed that words with low phonotactic probability seem to be learned more readily by children during the first three years of life, with a follow up study of nonword learning in 4-year-old children confirming this pattern (Storkel & Lee, 2011). Similar patterns are observed in adults, with a high-probability advantages shown for spoken nonword recognition (Luce & Large, 2001; Vitevitch & Luce, 1999), nonword repetition (Vitevitch & Luce, 1998, 2005), and serial nonword recall (Thorn & Frankish, 2005), but a disadvantage for high probability items in word learning (Storkel et al., 2006). This contrasting effect of phonotactic probability on word learning in both children and adults is hypothesized to be due to low probability sequences of speech sounds being more easily identified as novel words that need to be learned, effectively triggering the learning process more readily (Storkel, 2009; Storkel et al., 2006; Storkel & Lee, 2011).

Like any auditory signal, the speech signal consists of fluctuations in air pressure over time. The timing of these fluctuations heavily influences the perceptual outcome. Thus, speech processing is inherently tied to efficient encoding of the temporal structure of the speech signal (Kotz & Schwartze, 2010). On short timescales, variations in timing of formant transitions and voice-onset time determine

the perceived phoneme, while variations of the amplitude envelope at longer timescales defines the syllable rate, as well as metric and prosodic cues relevant for the processing of speech rhythm. Temporal regularities in the speech signal enable anticipation of the timing of upcoming events. While not precisely periodic, the average speech rate across languages lies between 3 and 8 syllables per second, overlapping with frequency ranges represented in the delta/theta range which are typically associated with tracking the amplitude envelope of the speech signal (e.g., Giraud & Poeppel, 2012). This allows us to direct our attention and neural resources towards the anticipated timing of upcoming speech events (e.g., syllables) to process the speech signal more efficiently. In addition to the syllable rate, the speech signal also has characteristic rhythmic properties that arise from the alternation of strong and weak, or stressed and unstressed syllables. These rhythmic properties can guide the parsing of the speech signal into words (Jusczyk et al., 1999; Thiessen & Saffran, 2003), and in some cases disambiguate the word (e.g. *pre*sent – pre*sent;* Cutler & Van Donselaar, 2001). The rules of these rhythmic regularities are specific to the language at hand, where certain patterns may be more common than others. For example, bisyllabic Dutch words are predominantly stressed on the first syllable.

Motor predictions are formulated during speech production to efficiently monitor and adjust our speech output. Several models have been developed to explain speech monitoring processes, combining frameworks from linguistic theory, perception and motor control (for recent review, see Gauvin & Hartsuiker, 2020). A key property of monitoring frameworks is the concept of an internal loop, analogous to a prediction of the sensory consequences of a motor plan, which allows the speaker to detect and correct speech errors at a faster rate than if they relied on the classical external speech perception pathways alone (Levelt, 1983; Nooteboom & Quené, 2017). Some models of speech monitoring can be considered direct extensions of the forward model, including elements such as the efference copy used to generate predictions of the sensory consequences to then be compared to the actual auditory feedback (Hickok, 2012; Pickering & Garrod, 2013). While these models are limited in their applications to lexical, conceptual, or syntactic monitoring (Gauvin & Hartsuiker, 2020), they account for a range of behavioral and neural observations of compensation for disrupted auditory feedback. When speaking in a noisy environment, we often subconsciously adjust the loudness of our voice and might find

ourselves inadvertently shouting, a phenomenon known as the Lombard effect. Similarly, if the auditory feedback of our own voice is manipulated, e.g. through shifts in pitch or formants, our articulatory system will adjust its output to correct for this (Katseff et al., 2012; Larson et al., 2007). Such corrections can also occur following self-produced deviations from the prototype of a target utterance (Niziolek et al., 2013). Regularities in the formal and temporal structure of utterances may also influence their production. In nonword repetition tasks, items with high phonotactic probability are repeated faster and more accurately than those with low phonotactic probability, in children, adolescents and adults (Edwards et al., 2004; Munson, Edwards, et al., 2005; Munson, Kurtz, et al., 2005; Vitevitch & Luce, 1998, 2005).

During development, we acquire implicit knowledge of the regularities in the formal and temporal structure of our native language(s). Sensitivity to these regularities appears to contribute to efficient perception and may additionally be exploited during reading development. Developmental dyslexia is characterized as persistent reading difficulties despite adequate educational opportunities (Lyon et al., 2003). While prevalence rates vary greatly due to differences in orthographic transparency across languages, educational systems and diagnostic criteria across countries and institutions, most estimates suggest that up to 10% of children are affected by dyslexia (Wagner et al., 2020). Aside from the characteristic reading disfluency, dyslexic readers typically show difficulties in tasks related to sublexical speech processing (Saksida et al., 2016)These difficulties may involve atypical categorical perception (Noordenbos et al., 2012; Serniclaes et al., 2004; but see e.g., Romanovska et al., 2019), or impaired lexical stress perception (Goswami et al., 2013; Jiménez-Fernández et al., 2015; Leong et al., 2011). It has been proposed that individuals with dyslexia have impaired statistical learning abilities (Gabay et al., 2015; Zhang et al., 2021; but see Schmalz et al., 2017), which would render them with reduced sensitivity to regularities in the speech signal, such as phonotactic or transitional probabilities. Similarly, some evidence for reduced encoding of amplitude rise times in the speech signal suggests atypical rhythmic capabilities in dyslexia (Goswami, 2011; Holliman et al., 2010; Ladányi et al., 2020). However, there is considerable heterogeneity in the sensory or perceptual deficits associated with dyslexia, which, among others, may be related to cross-linguistic variability in phonological structure, where auditory cues at different timescales may differ in their

relevance (Lallier et al., 2017). Thus, it is of interest to investigate the relationship between sensitivity to phonotactic and rhythmic properties of speech in individuals with dyslexia, to understand the relationship of these features with successful reading development.

3   Electrophysiological correlates of predictive processing

A variety of neural correlates have been employed to investigate predictive processing in the brain. Electroencephalography (EEG) is a particularly useful tool for investigating the underlying temporal dynamics given its high temporal resolution. EEG is recorded non-invasively via electrodes that are placed on the scalp, typically affixed in a cap. The signal picked up by the electrodes consists of a summation of postsynaptic potentials originating from groups of neurons oriented in parallel fashion. While EEG can pick up signals originating from both tangential and radial dipoles, its sensitivity is higher for radial sources located at the surface of the brain. Crucially, for a signal to be large enough to be detected non-invasively, large ensembles of neurons must fire synchronously. For an overview of methodological details of EEG, the reader is referred to works by Cohen (2014) and Luck (2014).

A central tenet in frameworks on predictive processing is that the anticipation of upcoming events allows us to process these more efficiently, as only the deviation from the expected event, or the prediction error, needs to be relayed to higher processing areas, resulting in less activation for predictable events. Such patterns are observed in several event-related potentials (ERPs), ranging from early to late time windows. Given the focus of this dissertation on sublexical regularities, this section will discuss early ERP correlates of this processing level, specifically the N1, P2, and mismatch negativity (MMN). To further elucidate underlying neural dynamics, this dissertation also employs time-frequency analysis. Thus, this section will further present oscillatory mechanisms associated with predictive processing, as well as time-frequency modulations observed in the paradigms applied in the current dissertation.

3.1   Event-related potentials

Analysis of event-related potentials (ERPs) involves averaging the signal across many trials. This approach highlights changes in voltage that are time- and phase-locked to

a critical event. ERP components measured at the scalp can be characterized by their polarity (positive vs negative), amplitude, latency, and topographic distribution. The ERP technique is particularly useful due to its high temporal precision and accuracy, and provides a relatively simple and intuitive approach to investigating differences in neural activity across conditions. A further advantage is that a vast body of literature from decades-long research characterizing the functional significance of ERP components can be used to interpret and contextualize results.

The N1 and P2 belong to the group of obligatory cortical auditory evoked potentials, meaning that their presence and morphology (latency and amplitude) are determined primarily by the acoustic features of the auditory input and the integrity of the ascending auditory pathway (Cone-Wesson & Wunderlich, 2003; Purdy et al., 2001). The N1, or N100, is characterized by a negative deflection at frontocentral electrode sites, peaking around 90 - 160 ms after the onset of an auditory stimulus. While frequently studied as a unitary event, the N1 deflection observed in the average ERP waveform is thought to be composed of several temporally overlapping subcomponents representing distinct neural processes that originate from separate cerebral generators (Näätänen & Picton, 1987). Some of these subcomponents, generated in primary and secondary auditory cortices, reflect specific acoustic properties of the stimulus. This is reflected in observations of changes in amplitude and latency of the N1 in response to changes in stimulus intensity (e.g., Picton et al., 1976), frequency (e.g., Picton et al., 1978), or slope of the energy change (e.g., Onishi & Davis, 1968). In addition to these specific auditory functions, a further subcomponent of the N1 is thought to reflect an "unspecific" transient arousal response or attention-triggering process in response to potentially relevant sensory input from various modalities (Näätänen & Picton, 1987).

The N1 is often studied in combination with the P2, or P200. The P2 is characterized as a positive deflection at central electrode sites occurring at approximately 150 – 250 ms following an auditory stimulus. Historically, it has often been studied together with the N1 in a peak-to-peak measurement, however, recent evidence suggests that these deflections represent functionally dissociable underlying processes (Crowley & Colrain, 2004). Similarly to the N1, P2 amplitude and latency are modulated by acoustic features of the stimulus such as intensity (Adler & Adler, 1989) and frequency (Wunderlich & Cone-Wesson, 2001), however

the nature of the modulations differ. The precise functional significance of the P2 is not well characterized. Positivities in the P2 latency window are evoked in a number of experimental paradigms, however, a likely role seems to be in attention and stimulus classification (García-Larrea et al., 1992). Convergent evidence suggests the P2 may be generated by several neural sources that are dissociable from N1 generators (Crowley & Colrain, 2004).

In addition to being modulated by low-level acoustic features of auditory stimuli, the N1 and P2 components have also been shown to be modulated by some linguistic features. For example, attention to speech sounds has been associated with a left-lateralization of the M100 (neuromagnetic equivalent of the N1; Parviainen et al., 2005; Poeppel et al., 1996). Furthermore, the N1-P2 amplitude can reflect phonetic cues that are relevant to speech perception such as amplitude rise time and formant transitions (Carpenter & Shahin, 2013), and is later and reduced for syllables presented in background noise compared to quiet (Koerner & Zhang, 2015). The amplitude of the P2 component specifically has been implicated as a marker of perceptual discriminability of speech sounds (Sheehan et al., 2005) and categorical perception of phonemes (Bidelman et al., 2013, 2020) and may serve an additional function in cohort reduction during visual gating of written words (Bles et al., 2007).

N1 and P2 are further modulated by the predictability of an event. Evidence from this comes from paradigms investigating repetition suppression (sometimes termed repetition positivity) and motor-induced suppression of the N1 components. In isochronous stimulus sequences, N1 amplitude decreases with increasing number of repetitions (Costa-Faidella et al., 2011), which has been linked to increased prediction accuracy (Auksztulewicz & Friston, 2016; Baldeweg, 2006; Friston, 2005; Winkler et al., 2009). In contrast, the P2 component is typically thought to be relatively spared from habituation processes (Kenemans et al., 1989). A large body of literature has investigated the N1 and P2 in the context of motor predictions, where self-initiated and externally generated stimuli are contrasted with each other. Here, self-initiated stimuli show reduced N1 and P2 amplitudes compared to the identical stimulus when it is externally triggered (e.g., Aliu et al., 2009; Bäss et al., 2008; Knolle et al., 2012, 2013a, 2013b, 2019; Niziolek et al., 2013; Ott & Jäncke, 2013; Pinheiro et al., 2018; SanMiguel et al., 2013; Sowman et al., 2012), a phenomenon referred to as motor-induced suppression. The degree of this suppression reflects the accuracy of

the prediction: more accurate predictions elicit greater suppression (e.g., Bäss et al., 2008; Niziolek et al., 2013). The suppression patterns may differ for different N1 subcomponents (SanMiguel et al., 2013). While attention also influences N1 amplitudes, this is likely independent of the effects of prediction (Timm et al., 2013; Foldal et al., 2020).

This suppression phenomenon has also been studied in the context of speech production. Some studies applying altered auditory feedback of utterances (e.g. formant or pitch shifts) report alterations in suppression, where unaltered feedback, or predictable feedback manipulations are associated with greater suppression (Heinks-Maldonado et al., 2005; Scheerer & Jones, 2014). Interestingly, these suppression modulations are distinct for N1 and P2 components, suggesting different underlying prediction processes (Sato & Shiller, 2018). Even in unaltered voice feedback, modulations of the suppression effect are observed, depending on whether a specific utterance is comparable to a prototypical utterance (Niziolek et al., 2013) or previous utterances of the same sound (Sitek et al., 2013), suggesting that the suppression mechanism is tuned for each specific production.

Another component often studied in the context of predictive processing is the mismatch negativity (MMN). This component is evoked in an oddball paradigm, where participants are presented with a sequence of stimuli consisting of a frequently occurring standard stimulus and one or more less frequent deviant stimuli. Here, the deviant stimuli elicit an increased negativity at frontocentral electrode sites around 100 – 250 ms after the onset of the deviation. The MMN is interpreted as a marker for automatic auditory change detection, and is elicited independently of task demands (for review, see Näätänen et al., 2007). Crucially for the aim of the current dissertation, the sensitivity of the MMN to stimulus features is experience dependent, meaning it can be thought of as a marker of learning-induced plasticity. In the study of language, it can be used to investigate the listener's sensitivity to certain linguistic features of the stimuli. Here, a larger MMN amplitude or shorter MMN latency are typically interpreted as indicators of facilitated change detection, suggesting more robust sensory memory traces. For example, the MMN is sensitive to categorical phoneme perception, with deviants crossing phoneme boundaries eliciting a larger MMN than those within boundaries (Aaltonen et al., 1987; Dehaene-Lambertz, 1997; Phillips et al., 2000; but see Ylinen et al., 2006). This pattern emerges during early

development, as infants learn their native language's phoneme categories (Cheour et al., 1998).

Previous work has demonstrated a sensitivity of the MMN to phonotactic probability (Bonte et al., 2005, 2007; Noordenbos et al., 2013). When using legal phonotactic structures, the pattern of results suggests facilitated change detection, with larger MMN amplitudes for high phonotactic probability deviants. Interestingly, this pattern is reversed when deviants violate phonotactic constraints of the language (Steinberg et al., 2011). Several studies have also investigated the sensitivity of the MMN to variations in lexical stress. Many of these studies have been conducted in languages with fixed stress patterns, such as Finnish (Ylinen et al., 2009), or Hungarian (Honbolygó et al., 2004; Honbolygó & Csépe, 2013; Ragó et al., 2014). In these languages, stress always lies on the first syllable, thus changes in this pattern violate the constraints of the language. In line with violations of phonotactic constraints, MMN modulation by variations in stress patterns appear to follow a violation response: when comparing identical stimuli presented as deviant and standard to each other, deviants with illegal stress patterns elicit a larger MMN than those with legal stress (Honbolygó & Csépe, 2013). In languages with variable stress patterns such as Dutch, English or German, studies of MMN sensitivity to variations in stress have primarily focused on processing of the perceptual markers of lexical stress (pitch, intensity and syllable duration; Tong et al., 2014; Zora et al., 2015), or have not directly compared the magnitude of the MMN to different stress patterns, but simply noted its presence or absence (Weber et al., 2004) or differences between groups (Schaadt & Männel, 2019). Additionally, ERP correlates of phonotactic probability and syllable stress have traditionally been studied in isolation, while they vary simultaneously in natural speech.

The current dissertation investigates the neural sensitivity to variations in phonotactic probability and syllable stress regularities when passively presented and self-generated, as indexed by the N1, P2 and MMN. In Chapter 2, we develop a multi-feature passive oddball paradigm to investigate processing of phonotactic probability and syllable stress simultaneously, filling the described gaps in the literature. In Chapter 3, we extend this approach to testing adult dyslexic readers. In Chapter 4, we investigate the sensitivity of the N1 and P2 components to phonotactic probability and syllable stress in self- and externally triggered speech.

1.3.2 Neural oscillations

While ERPs can give simple and intuitive insights into predictive processing, with overall patterns suggesting reduced neural activation, i.e., more efficient processing for predictable events, they do not tell us much about the underlying mechanisms that lead to these different outcomes. To gain further insights into how the brain processes critical events, it can be useful to investigate the time-frequency domain. Cortical oscillations have been associated with functions in several cognitive domains, and provide a mechanism for flexible and dynamic communication between areas (Fries, 2005). When quantifying the properties of an oscillation, we can determine its frequency, phase, and power. Cortical oscillations are typically grouped into frequency bands: delta = ~1 - 4 Hz, theta = ~4 - 8 Hz, alpha = ~8 - 12 Hz, beta = ~ 12 - 30 Hz, gamma = > 30 Hz. The functional interpretation of modulations in these frequency bands can differ depending on the task or cognitive domain being studied. In the following section we provide an overview of oscillatory correlates of prediction, speech and language processing, as well as auditory change detection associated with the mismatch negativity.

Cortical oscillations have also been implicated in predictive processing (Engel et al., 2001; Morillon & Schroeder, 2015). In the context of formal and temporal predictions (Arnal & Giraud, 2012), dissociable cortical oscillatory mechanisms have been proposed. A central process underlying temporal prediction is thought to be a phase reset in slow delta-theta oscillations to align the "ideal phase" of the oscillation with anticipated input for efficient processing (Schroeder & Lakatos, 2009). This phase-reset has also been observed with temporally predictable yet arhythmic stimuli (Breska & Deouell, 2017). Temporally predictable events have also been associated with a reduction in alpha power (Rohenkohl & Nobre, 2011), signaling a decrease in active inhibition, or conversely increased attention towards an event (Jensen et al., 2012; Klimesch, 2012). Finally, beta power is coupled with delta/theta phase to track temporal regularities in a signal (Arnal et al., 2015; Cravo et al., 2011). Given the established role of beta oscillations in motor activity, this highlights once more the close link between temporal prediction and the motor system (Kotz & Schwartze, 2011; Morillon & Schroeder, 2015). This beta modulation is thought to serve the function of synchronizing relevant neuronal populations for anticipated sensory

input, thereby also contributing to formal prediction mechanisms (Arnal & Giraud, 2012; Bressler & Richter, 2015; van Pelt et al., 2016). The magnitude of the prediction error, or mismatch between expected and actual sensory input, is represented by gamma power (Brodski et al., 2015; van Pelt et al., 2016). This proposed oscillatory framework highlights the notion that formal and temporal predictions are generated via distinct yet interacting processes.

Cortical oscillations have also been attributed specific functions in speech processing (Giraud & Poeppel, 2012; Poeppel, 2003). A key feature of the speech signal is that units of analysis are organized at different time scales. While the speech signal is not precisely periodic, it is nevertheless characterized by temporal regularities. Rapid spectral transitions at the scale of 30-50 Hz mark phonemic information such as formant transitions or voice-onset times, while slower temporal modulations around ~3-8 Hz of the amplitude envelope relate to the timing of syllables. Finally, prosodic information indicative of the timing of words and phrases occurs at rates around 1-2 Hz (Poeppel. 2003). Following these temporal modulations at different time scales, it has been proposed that auditory cortical oscillations allow segmentation of the speech signal into the relevant linguistic units, where low frequency delta/theta oscillations track the amplitude envelope of incoming speech, and through theta-gamma nesting align neuronal excitability in the low gamma range (~30-50 Hz) to the syllable rate for fine-grained analysis at the phonemic level (Rimmele et al., 2021). At higher linguistic processing levels, theta and alpha oscillations have been linked to lexical retrieval (Bastiaansen et al., 2005; Mellem et al., 2013; Peña & Melloni, 2012), while beta and gamma oscillations have been linked to syntactic and semantic processing, respectively (Bastiaansen et al., 2010; Bastiaansen & Hagoort, 2015; Peña & Melloni, 2012).

While the ERP correlates of auditory change detection as studied in the oddball paradigm have been well characterized, recent directions have begun to approach this process from an oscillatory perspective to elucidate underlying neural mechanisms. Given the vast analytical flexibility in the time-frequency domain, synthesizing the results of these investigations into a cohesive picture is challenging, however some patterns emerge that mirror observations in the ERP domain and contribute to their functional interpretation. A dominant observation in time-frequency analyses auditory change detection is an increase in theta ITC for the deviant relative to the

standard (e.g., Fuentemilla et al., 2008; Hsiao et al., 2009), sometimes also associated with an increase in theta power. These theta modulations have been shown to be associated with perceptual discrimination abilities (Bishop et al., 2011; Jin et al., 2014) and speech intelligibility (Koerner et al., 2017). Additionally, the theta modulations may differ depending on language abilities (Cantiani et al., 2019; Halliday et al., 2014). Chapter 3 of the current dissertation investigates the time-frequency correlates of auditory change detection in the passive oddball paradigm. Specifically, the aim was to understand whether formal and temporal deviants are associated with distinct neural mechanisms that can be attributed to formal and temporal prediction, with a focus on low-frequency oscillations as these are associated with tracking phonotactic (Di Liberto et al., 2019) and syllable stress regularities, and are hypothesized to be impaired in dyslexic readers (Goswami, 2011; Lallier et al., 2017).

4   Overview

This doctoral dissertation presents the findings from three EEG experiments examining the role of sublexical formal and temporal regularities in speech processing. In Chapter 2, a multi-feature passive oddball paradigm with stimuli consisting of Dutch pseudowords is developed, to test the sensitivity of the mismatch negativity to variations in phonotactic probability and syllable stress in Dutch pseudowords. Chapter 3 extends this approach to investigate the sensitivity of these features in individuals with dyslexia, investigating the modulation of the MMN and the associated time-frequency correlates. Chapter 4 employs a motor-to-auditory paradigm to investigate the sensitivity of the N1 and P2 components to variations in phonotactic probability and syllable stress.

Chapter 2

ERP mismatch response to phonological and temporal regularities in speech

Based on

Emmendorfer, A. K., Correia, J. M., Jansma, B. M., Kotz, S. A., & Bonte, M. (2020). ERP mismatch response to phonological and temporal regularities in speech. *Scientific reports*, *10*(1), 1-12.

ABSTRACT

Predictions of our sensory environment facilitate perception across domains. During speech perception, formal and temporal predictions may be made for phonotactic probability and syllable stress patterns, respectively, contributing to the efficient processing of speech input. The current experiment employed a passive EEG oddball paradigm to probe the neurophysiological processes underlying temporal and formal predictions simultaneously. The component of interest, the mismatch negativity (MMN), is considered a marker for experience-dependent change detection, where its timing and amplitude are indicative of the perceptual system's sensitivity to presented stimuli. We hypothesized that more predictable stimuli (i.e. high phonotactic probability and first syllable stress) would facilitate change detection, indexed by shorter peak latencies or greater peak amplitudes of the MMN. This hypothesis was confirmed for phonotactic probability: high phonotactic probability deviants elicited an earlier MMN than low phonotactic probability deviants. We do not observe a significant modulation of the MMN to variations in syllable stress. Our findings confirm that speech perception is shaped by formal and temporal predictability. This paradigm may be useful to investigate the contribution of implicit processing of statistical regularities during (a)typical language development.

1   Introduction

In order to effectively deploy resources for efficient processing of incoming sensations from our environment, our brain formulates online predictions of upcoming sensory events (Engel et al., 2001; Friston, 2012; Raichle, 2010). This is possible through our knowledge about regularities in the sensory environment, allowing us to anticipate the consequences of an action, adapt behaviour to an upcoming event, or ease sensory processing under noisy conditions. These predictions may be formal ('what') or temporal ('when') in nature (Arnal & Giraud, 2012). A formal prediction constitutes a prediction of the formal structure or content of an upcoming event. In speech processing, formal predictions can occur at multiple levels, such as the semantic category of a word in a sentence, or the sequence of speech sounds (phonemes) within words, the phonotactic structure. Temporal predictions on the other hand are related to the anticipation of temporally regular events. Within language, the vocalic nucleus of a syllable is often considered the perceptual beat (Port, 2003), and the metre (alternation between strong and weak beats) can be described as the alternation between strong and weak (or stressed and unstressed) syllables. Therefore, the anticipation of syllable stress may constitute a form of temporal prediction.

Formal and temporal predictions are thought to operate via distinct neural oscillatory mechanisms (Arnal & Giraud, 2012) and structural circuits (Kotz & Schwartze, 2010, 2016). While the processing of formal and temporal structure of the speech signal has traditionally been studied in isolation, these features vary simultaneously in natural speech. A few studies have found variation in metre to influence semantic (Rothermich et al., 2012) and syntactic (Schmidt-Kassow & Kotz, 2009) processing. Furthermore, difficulties in processing the temporal structure of speech have been suggested to underlie phonological processing deficits observed in dyslexia (Goswami, 2011; Ladányi et al., 2020; Lallier et al., 2017). It is therefore of interest to study how formal and temporal predictability may interactively influence speech perception.

Both forms of predictions are established through experience in development and may play a vital role in successful skill learning. Evidence of sensitivity to regularities of the formal and temporal structure of language can already be found in infants. Newborns within 5 days of birth are already sensitive to differences in the

rhythmic structure of speech (Nazzi et al., 1998), while sensitivity to statistical regularities between neighbouring speech sounds has been demonstrated as early as 8 months of age (Saffran et al., 1996). Sensitivity to syllable stress and phonotactic probability provide a crucial foundation for early language development, allowing infants to segment words in the continuous speech signal (Jusczyk et al., 1999; Mattys & Jusczyk, 2001; Thiessen & Saffran, 2003).

Phonotactic probability continues to influence performance on a number of primarily sublexical language processes throughout the lifespan. Children and adolescents show better performance in speed and accuracy for high compared to low phonotactic probability items in nonword repetition tasks (Edwards et al., 2004; Munson, Edwards, et al., 2005; Munson, Kurtz, et al., 2005). This effect is reversed for word learning (Storkel, 2009; Storkel & Lee, 2011). Similar patterns are observed in adults, with a high-probability advantages shown for spoken nonword recognition (Luce & Large, 2001; Vitevitch & Luce, 1999), nonword repetition (Vitevitch & Luce, 1998, 2005) and serial nonword recall (Thorn & Frankish, 2005), but a disadvantage for high probability items in word learning (Storkel et al., 2006). This contrast between nonword repetition and word learning in both children and adults is hypothesized to be due to low probability sequences of speech sounds being more easily identified as novel words that need to be learned, effectively triggering the learning process more readily (Storkel, 2009; Storkel et al., 2006; Storkel & Lee, 2011).

Although these and other behavioural effects of phonotactic probability are relatively well documented (see review by Auer & Luce, 2005)), the role of lexical stress in speech perception, beyond guiding speech segmentation, is less well studied. When a language permits different lexical stress patterns, these may guide the resolution of lexical conflict in spoken word recognition (Cutler, 2005). Performance on nonword repetition has also been shown to improve for more "typical" stress patterns within the language (Vitevitch et al., 1997).

We aim to probe neural correlates of these processes in normally reading adults by means of a passive oddball paradigm, which is particularly suited for the investigation of experience-dependent neurophysiological changes. In a classical passive oddball paradigm, a sequence of auditory stimuli is presented to the participant, consisting of a frequently occurring standard stimulus, and an infrequent

deviant or 'oddball' stimulus. The participant is instructed to ignore the stimuli and typically reads a book or watches a silent film to remain awake and relaxed. In this type of passive paradigm, the ERP component of interest is the mismatch negativity (MMN), a negative deflection in voltage surrounding frontocentral and central electrodes in the window 100 – 250 ms after the onset of the stimulus deviation (Näätänen et al., 2007). While early MMN studies have used simple stimuli such as pure tones (Näätänen et al., 1978), the component has also been employed to study linguistic processing, ranging from simple vowel discrimination (Aaltonen et al., 1987) to higher-order processes such as syntax (Pulvermüller & Shtyrov, 2003). The MMN has been interpreted as a marker for experience-dependent change detection and its timing and amplitude are indicative of the perceptual system's sensitivity to the presented stimuli.

Oddball experiments in adults and children have demonstrated that the MMN component can be modulated by variations in phonotactic probability, i.e. the probability of the co-occurrence of phonemes in a language, where deviants with higher probability have been shown to elicit larger mismatch responses compared to deviants with low probability (Bonte et al., 2005, 2007; Noordenbos et al., 2013). The paradigm has also been applied to study processing of syllable stress patterns in both real and pseudowords (Honbolygó et al., 2004; Honbolygó & Csépe, 2013; Ragó et al., 2014; Ylinen et al., 2009). In languages with a strict syllable stress pattern such as Hungarian and Finnish, where stress is always on the first syllable in bisyllabic words, deviant stimuli using an illegal stress pattern elicit two consecutive MMNs, where the first is thought to reflect the missing stress on the first syllable, while the second reflects the detection of the unexpected stress on the second syllable (Honbolygó et al., 2004; Honbolygó & Csépe, 2013; Ylinen et al., 2009). Although variations of formal and temporal predictability simultaneously occur in natural speech, ERP studies have so far typically investigated these in isolation.

In the current study, we employed a multi-feature oddball paradigm simultaneously manipulating both formal and temporal predictions in Dutch pseudowords, in the form of phonotactic probability and syllable stress pattern, respectively. We examined the effect of violations of these predictions on the MMN response, where we expected the timing and magnitude of this response to vary with the formal and temporal predictability of the stimuli. If more predictable formal and

temporal features of the stimuli (i.e. high phonotactic probability and first syllable stress) are processed more efficiently, this should lead to easier change detection, indexed by greater MMN peak amplitudes and/or shorter latencies. This has been shown for phonotactic probability (Bonte et al., 2005, 2007; Noordenbos et al., 2013). However previous studies investigating MMN sensitivity to syllable stress have been primarily conducted in languages with fixed-stress patterns (Honbolygó et al., 2004; Honbolygó & Csépe, 2013; Ragó et al., 2014; Ylinen et al., 2009). Those conducted in languages with variable stress patterns (e.g. English, German), primarily focussed on MMN sensitivity to specific acoustic features of syllable stress, comparing responses between naturally spoken first syllable stress standards to deviants where pitch, intensity or vowel duration is manipulated to generate second syllable stress (Tong et al., 2014; Zora et al., 2015), or simply note the presence of an MMN to both first and second syllable stress deviants without directly comparing the two (Weber et al., 2004). Additionally, to extend upon previous studies which investigated these features in isolation, we aimed to test whether their simultaneous manipulation would yield similar or different patterns of MMN modulations, and whether stimulus features would interactively modulate the mismatch response (i.e. whether variations in syllable stress modulate formal deviant processing and vice versa).

In summary, the current study aims to test the following hypotheses: (1) deviants differing from standards in terms of phonotactic probability (hereafter formal deviants) or syllable stress (hereafter temporal deviants) elicit an MMN, indicated by a greater negativity in response to deviants compared to standards; (2) this MMN to formal or temporal deviants is modulated by phonotactic probability or syllable stress, respectively, which may present as a larger MMN amplitude (Bonte et al., 2005, 2007; Noordenbos et al., 2013), or shorter MMN latency for more predictable (high phonotactic probability/first syllable stress) deviants. (3) variations in predictability in the other domain (*syllable stress* for formal deviants, *phonotactic probability* for temporal deviants) may further modulate this MMN sensitivity. The analyses compared identical stimuli presented in different conditions (standard versus formal or temporal deviant), which allowed us to generalize the results beyond mere acoustic differences between the stimuli.

2   Methods

2.1   Participants

29 native Dutch-speaking participants with normal reading skills participated in the experiment after giving their informed consent. 5 participants were excluded from further analysis (1 for technical issues during recording, 1 for excessive noise in EEG data (>20% trials rejected from amplitude criterion), 2 for exclusion criteria revealed during or after participation (1 left-handed, 1 learning disability), 1 for failure to complete both study visits), leaving a final sample of 24 right-handed participants (mean age = 22.6; range = 18 - 30, 10 males). The study was approved by the Ethics Committee of the Faculty of Psychology and Neuroscience at Maastricht University performed in accordance with the approved guidelines and the Declaration of Helsinki.
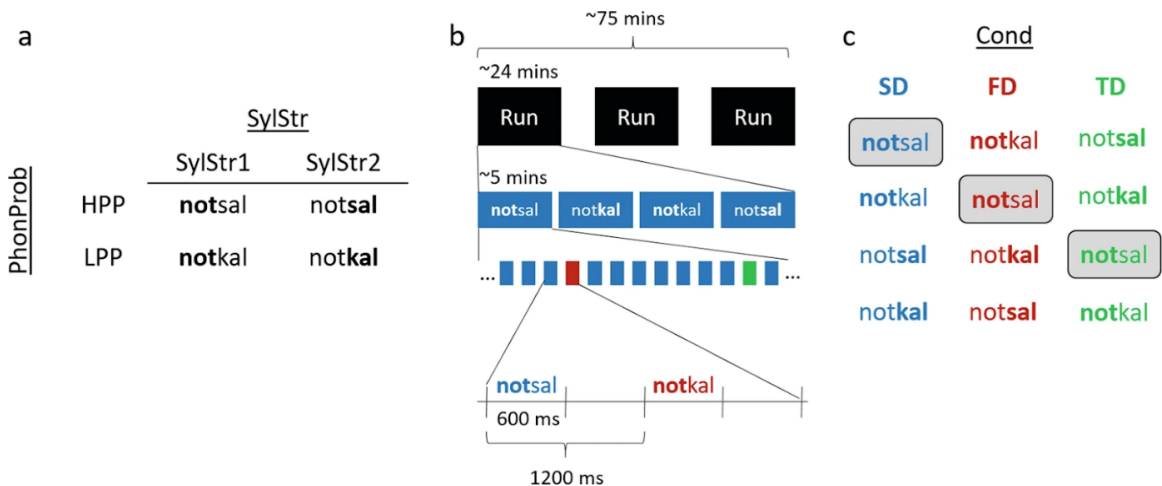
Figure 1: Experimental design. (a) Pseudoword stimuli varying in phonotactic probability (PhonProb) and syllable stress (SylStr). Bold font indicates stressed syllable (SylStr1 = first syllable, SylStr2 = second syllable. The phoneme combination '-ts-' constitutes high phonotactic probability (HPP) and '-tk-' low phonotactic probability (LPP). (b) Overview of experimental session. (c) Each stimulus is presented as standard (SD), formal deviant (FD), and temporal deviant (TD) in separate conditions (Cond), allowing the comparison of identical stimuli across conditions (example highlighted for **not**sal).

2.2   Stimuli

2.2.1   Pseudowords

The stimuli used in the oddball paradigm were adapted from a previous paradigm employing Dutch pseudowords *notsel* and *notkel* (Bonte et al., 2005, 2007). These stimuli were initially constructed by calculating phonotactic probabilities using the CELEX database (Baayen et al., 2001), where the phonotactic structure '-*ts*-' was found

to have a higher probability than '-*tk-*'. This stimulus pair can therefore be used to test the role of formal predictions. To add the dimension of temporal prediction to these stimuli, we additionally varied the syllable stress pattern placing the stress on either the first or second syllable, creating a stimulus quadruplet (Figure 1a). We adapted the pseudoword pairs from notsel-notkel to notsal-notkal in order to avoid possible changes of the vowel 'schwa' due to stress variation. Both phonotactic constructions and syllable stress patterns are legal in Dutch but occur at different frequencies. The relative frequencies of these features are indicated in Table 1, as determined by the word frequencies of bisyllabic Dutch words containing the target phoneme structure or syllable stress pattern, retrieved from the CELEX database (Baayen et al., 2001).

2.2.2   Recording and editing stimuli

The stimuli were spoken by a female native Dutch speaker and recorded in a sound attenuated chamber using GoldWave Digital Audio Editor (sampling rate 44100 Hz, 16 bit; GoldWave Inc., St. John's, NL Canada). Due to the scarcity of second syllable stress in bisyllabic words with a CVCCVC syllable structure (occurring only in 6% of CVCCVC words as indicated in Table 1), natural pronunciation of the pseudowords with second syllable stress can be challenging to Dutch speakers. To circumvent this issue, the speaker was instructed to pronounce the syllables of interest within the context of existing bisyllabic Dutch words, which contained the same (spoken) consonant cluster and stress pattern as the target pseudowords.

The speaker first pronounced the existing Dutch word several times to familiarize herself with it. The first or second syllable of the word was then replaced by the target syllable in the pseudoword, and the speaker was instructed to pronounce the new word with the same stress pattern as the original word. Thus, the speaker first pronounced the real word /**bad**zout/, followed by the pseudowords /**not**<u>zout</u>/ and /**bad**<u>sal</u>/ (bold font denotes syllable stress, underline denotes target syllable) to create our first syllable stress pseudoword /**not**sal/. The other syllables were constructed in a similar way: /ont**slag**/ -> /<u>not</u>**slag**/ & /ont**<u>sal</u>**/ -> /not**sal**; /**geld**kas/ -> /**not**kas/ & /**geld**<u>kal</u>/-> /**not**kal/ and /goed**koop**/ -> /<u>not</u>**koop**/ & /goed**<u>kal</u>**/ -> /not**kal**/. (Note that in Dutch, a syllable final /d/ is indistinguishable from a syllable final /t/ due to final-obstruent devoicing, and the /z/ in /badzout/ is pronounced as /s/). The target syllables were later spliced from these recordings and

34

combined to create the pseudowords using Praat (Boersma & Weenink, 2013). The matching consonant cluster at the syllable boundary ensured identical co-articulatory cues, facilitating cross-splicing of syllables to create the final pseudowords. To construct the pseudowords, excised target syllables were paired to ensure equivalent changes in pitch and intensity for both first or second syllable stress. *Notsal* stimuli were created by combining /no/ of the first syllable with /tsal/ from the second syllable to minimize acoustic artefacts within the consonant cluster /ts/ from the splicing procedures. Because of the voice-onset time preceding the /k/ in /-kal/, this was not necessary for *notkal* stimuli. The constructed stimuli were then edited to equalize for loudness (rms amplitude) and duration (600 ms). Three versions of each stimulus were created from distinct utterances of each syllable (i.e. each syllable in the final pseudowords was unique). This allowed the generalization to the target features of phonotactic probability and syllable stress beyond small acoustic variations.

Table 1

INL frequencies of target phoneme structures and syllable stress patterns

	PhonProb			**SylStr**		
	-ts- Freq	**-tk- Freq**	**-ts-/-tk- Ratio**	**SylStr1 Freq**	**SylStr2 Freq**	**SylStr1/SylStr2 Ratio**
All BiSyl	94061	7787	12.09	7633058	1926345	3.962
CVCCVC	1227	0	-	292794	18817	15.560

Note. Frequencies represent the sum of INL (Instituut voor de Nederlandse Taal) frequencies of all Dutch words containing the indicated consonant cluster or stress pattern, for all bisyllabic Dutch words (All BiSyl) or limited to those with a CVCCVC structure, while ratios represent the ratio of these frequencies for high probability (HPP, SylStr1) compared to low probability (LPP, SylStr2) conditions. (C = consonant, V = vowel, SylStr1 = 1st syllable stress, SylStr2 = 2nd syllable stress).

2.3   Data acquisition

2.3.1 Oddball paradigm

A passive oddball paradigm was used, where each stimulus served as the standard in separate conditions (Figure 1b,c). Each condition contained a temporal and a formal deviant, which differed from the standard in terms of either the syllable stress or phonotactic probability, respectively. Each condition contained a total of 1,620 trials (1,332 standards and 144 deviants, or 8.9%, per deviant type). The experiment took place over two sessions. Each session consisted of three runs of approx. 24 mins, split into four blocks (one per condition) of 270 trials. Participants were encouraged to take breaks as needed in between blocks and runs. Within a block, trials were presented with trial duration of 1,200 ms (i.e. inter-stimulus-interval 600 ms). The stimuli were presented in pseudorandom order, with deviants separated by 1-8 standards. The order of blocks within each run was randomized for each participant.

2.3.2 EEG recording

EEG was recorded with BrainVision Recorder (Brain Products, Munich, Germany) using a 63-channel recording setup. Ag/AgCl sintered electrodes were mounted in an EasyCap electrode cap (EASYCAP GmbH, Herrsching, Germany) according to the 10% equidistant system, including 57 scalp electrodes, left and right mastoids for offline re-referencing, and four EOG electrodes to facilitate removal of artefacts caused by eye movements. The skin at electrode sites was prepared with NuPrep Skin Prep Gel (DO Weaver and Co., USA) and an electrolyte gel was used to keep impedances below 10kΩ. Data were recorded at a sampling rate of 1000Hz, using Fpz as an online reference and AFz as ground. During recording, participants were seated on a comfortable chair in an acoustically and electrically shielded room and instructed to watch a silent nature documentary while ignoring the auditory stimuli.

2.4   Analysis

2.4.1 Preprocessing

Preprocessing was performed using MATLAB 2017a and the EEG analysis toolbox Letswave 6 (https://github.com/NOCIONS/letswave6). Data were first filtered (band pass 0.5 – 70 Hz, notch filter 48 – 52 Hz) and down-sampled to 250 Hz. Noise from eye-movements, muscle artefacts, and noisy electrodes was removed using

independent component analysis(Delorme et al., 2007) (ICA) with the runica algorithm implemented in Letswave 6, decomposing the signal into 63 components. The time course recorded during the breaks between blocks was removed from the data to reduce noise prior to the ICA. Artefactual components were selected for removal based on the time course and topography. A median of 19 components (~30%) was rejected per dataset (SD = 6). From the resulting data, -100 to 1000 ms epochs relative to the onset of the stimulus were extracted. After DC removal and baseline correction to the pre-stimulus interval (-100 to 0 ms), an automatic artefact rejection algorithm was applied with an amplitude criterion of 75µV over scalp electrodes to remove trials with remaining artefacts, and the data was re-referenced to the average mastoids. Deviants occurring after only one standard were excluded from analysis. Standards immediately preceding deviants were selected for analysis, resulting in up to 126 trials for each deviant type, and 252 standards for each stimulus. To allow comparing the same standard trials to both formal and temporal deviants, while ensuring equal number of trials across conditions, a random subset of standards was selected to equal the smallest number of deviants across conditions per participant. Within participants, the number of trials was equalized across conditions, leading to a final number of 99 – 124 trials per condition per participant.

2.4.2 ERP analysis

Trials were averaged after time-locking to the onset of the auditory deviation, corresponding to stimulus onset for temporal deviants, and the /t/-onset for formal deviants. Difference waves were calculated (deviant – standard of identical stimuli, where the standard was always time-locked to the same moment as the respective deviant). Individual and grand average difference waveforms per condition where examined to determine the time window for peak extraction (100 - 300 ms after /t/-onset for formal deviants, 200 - 350 ms after stimulus onset for temporal deviants). MMN peak latency to formal and temporal deviants was defined based on the difference waves per participant at FCz. This electrode was selected due to the well-documented frontocentral topography of the MMN (Näätänen et al., 2007). Amplitude measures of the MMN were determined around the FCz peak latency for all other electrodes to ensure the comparison of the same underlying process across the scalp (Luck, 2014). First an automatic algorithm in the Letswave toolbox was used to find a

negative peak at FCz in the pre-specified time window. The waveforms and topography were then visually inspected to confirm the selection, or to adjust it to a more fitting peak that reflected the typical frontocentral MMN topography, within a final time window of 80 - 300 ms for formal deviants, and 120 - 370 ms for temporal deviants. Mean amplitudes (+/- 24ms surrounding the peak) were calculated for standard and deviant waveforms at all other electrodes at this latency. Frontocentral (Fz, F1, F2, F3, F4, FCz, FC1, FC2, FC3, FC4) and centroparietal (Cz, C1, C2, C3, C4, CPz, CP1, CP2, CP3, CP4) regions of interest for the amplitude measures were selected based on the frontocentral topography of the elicited mismatch response and comparisons with previous literature(Bonte et al., 2005; Honbolygó & Csépe, 2013; Ylinen et al., 2009).

Statistical analyses were performed in R (version 3.6.3, R Core Team, 2013) with the Rstatix package (Kassambara, 2019). For formal and temporal deviants individually, the MMN mean amplitudes were analysed with a 2x2x2x2 repeated measures ANOVA with *phonotactic probability* (PhonProb: HPP vs. LPP), *syllable stress* (SylStr: SylStr1 vs. SylStr2), *condition* (Standard vs. Deviant) and *region-of-interest* (ROI: frontocentral vs. centroparietal) as within-subjects factors. We set out to test the following hypotheses: (1) formal and temporal deviants elicit an MMN, indicated by main effect of *condition* on MMN mean amplitude, with greater negativity for deviants compared to standards; (2) this MMN is sensitive to the predictability of the stimulus features (*phonotactic probability* x *condition* for formal deviants, *syllable stress* x *condition* for temporal deviants), with more predictable features (HPP or SylStr1) showing a greater mismatch response; (3) variations in predictability in the other domain (*syllable stress* for formal deviants, *phonotactic probability* for temporal deviants) may further modulate this MMN sensitivity (*phonotactic probability* x *syllable stress* x *condition).*

The MMN latency for both deviant types was further analyzed in a 2x2 repeated measures ANOVAs with PhonProb and SylStr as within-subjects factors (because the peak latency was determined based on the difference wave MMN peak at a single electrode, Cond and ROI were not included as factors in this analysis). Here we tested our hypotheses (2) the MMN latency is sensitive to the predictability of stimulus features, indicated by main effect of *phonotactic probability* for formal deviants and *syllable stress* for temporal deviants, with more predictable features

(HPP or SylStr1) showing an earlier mismatch response; and (3) variations in predictability in the other domain (*syllable stress* for formal deviants, *phonotactic probability* for temporal deviants) may further modulate this MMN sensitivity (*phonotactic probability* x *syllable stress*).

3   Results

Figure 2 provides an overview of the waveforms elicited by the stimuli presented within each condition (e.g. standard (SD) = **not**sal, formal deviant (FD) = **not**kal, temporal deviant (TD) = not**sal**), time-locked to the onset of the stimulus. Visual inspection of the waveforms revealed that all pseudoword contrasts elicited a mismatch response between 100 - 350 ms after the onset of auditory stimulus deviation. Formal deviants appeared to show a negative deflection compared to the standard in the window 100 - 300 ms after the auditory deviation at the /t/-onset (~350 – 550ms after stimulus onset), in all conditions but LPP SylStr2 (bottom right panel), while temporal deviants showed a similar negative deflection around 200 - 350 ms after stimulus onset. In order to focus on MMN modulations at a more abstract level of representation, in our further analyses we compared the activity elicited by the same stimuli as standards and deviants across blocks; e.g. *not*sal formal deviant minus *not*sal standard. The following results are presented for formal and temporal deviants separately, time-locked to the auditory deviation in the respective contrasts. Amplitude analyses were performed on mean amplitudes (+/- 24 ms surrounding peak). Here we present results of the tests of our a priori hypotheses, corrected for multiple comparisons using Bonferroni-Holm correction (Cramer et al., 2016). All other effects tested in the ANOVAs (significant and non-significant), can be found in the supplementary materials.
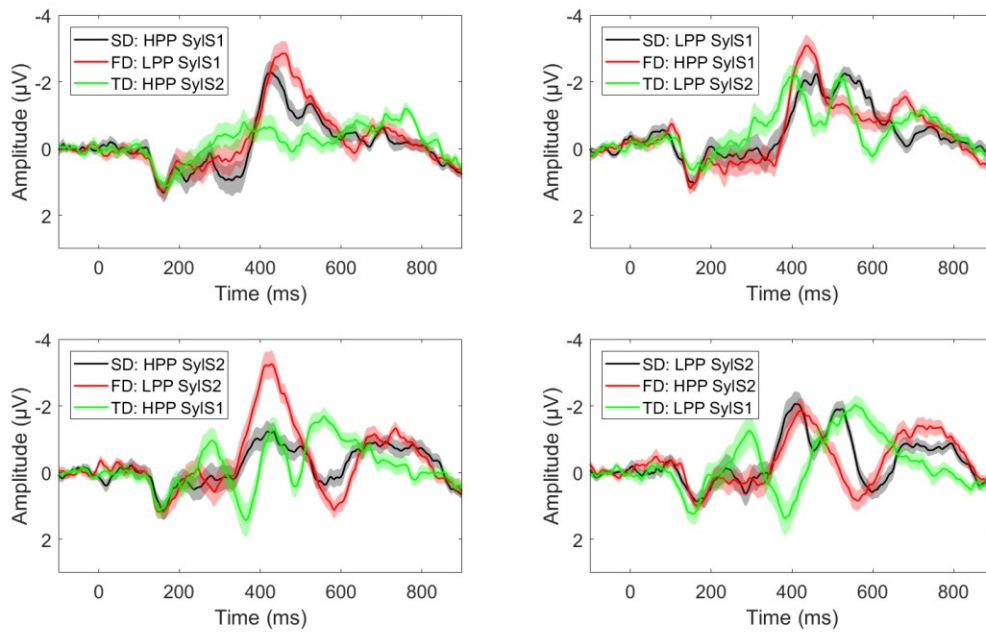
Figure 2: Grand average waveforms +/- SEM at a frontocentral ROI within conditions: Each panel represents one standard (black) and the formal (red) and temporal (green) deviants that were presented within a block: SD = Standard, FD = Formal deviant, TD = Temporal deviant.

3.1 Formal deviants

Grand average ERPs for standards and formal deviants, time-locked to the onset of the /t/ at the frontocentral ROI, are shown in Figure 3. Both HPP and LPP formal deviants, averaged across syllable stress, showed a more negative peak response compared to the identical stimulus presented as a standard in a window around 100 - 300 ms after the /t/-onset, with comparable topographies (Figure 4a). Visual inspection of the difference waves suggests an effect of phonotactic probability, with a larger mismatch response for LPP formal deviants and an earlier mismatch response for HPP deviants (top panel), but no effect of syllable stress on formal deviants (bottom panel).
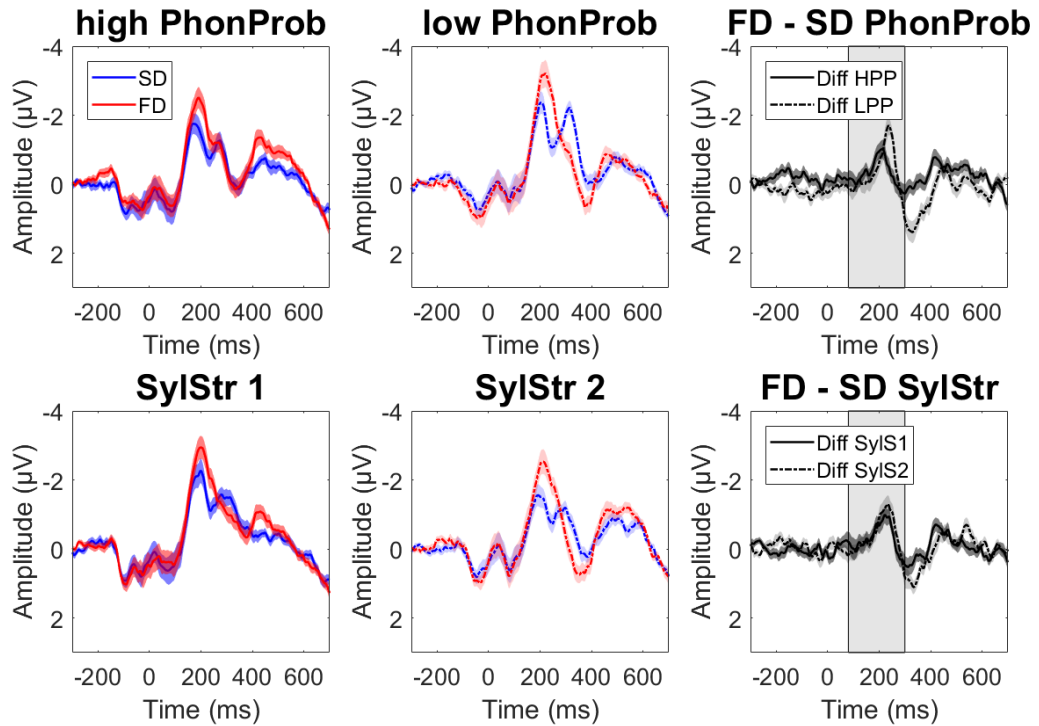
Figure 3: Formal deviants. Grand average waveforms +/- SEM at frontocentral ROI, time-locked to /t/-onset. Top panel is averaged across syllable stress: high phonotactic probability (red: deviant, blue: standard), low phonotactic probability, and the difference waves (dark grey: high PhonProb, light grey: low PhonProb). The bottom panel is averaged across phonotactic probability: first syllable stress, second syllable stress, and difference waves (dark grey: SylStr1, light grey: SylStr2). Shaded area in difference waves represents window for MMN peak extraction.

The amplitude observations were statistically tested using a 2x2x2x2 repeated measures ANOVA (PhonProb x SylStr x ROI x Cond) on MMN mean amplitudes. The analysis of mean amplitudes revealed a significant main effect of Cond [$F(1,23)$ = 107.642, p_{adj} < 0.001], with deviants eliciting a significant MMN, confirming our first hypothesis. The PhonProb x Cond interaction was not significant [$F(1,23)$ = 0.187, p_{adj} = 1.000], however, a possible trend towards a three-way PhonProb x Cond x ROI interaction [$F(1,23)$ = 8.568, p_{adj} = 0.112] is observed (Figure 5a). Post-hoc two-sided paired-samples t-tests on the mean amplitude difference (FD – SD) comparing high and low phonotactic probability, averaged across syllable stress, did not reveal any significant effect in either the frontocentral [$t(23)$ = 1.08, p_{adj} = 0.584] or centroparietal ROIs [t = -0.373, p_{adj} = 1.000] (while we hypothesized a larger MMN for HPP deviants, which would warrant a one-sided test, our data as seen in Figure 3 suggest a larger MMN or LPP deviants, therefore we selected the two-sided test).
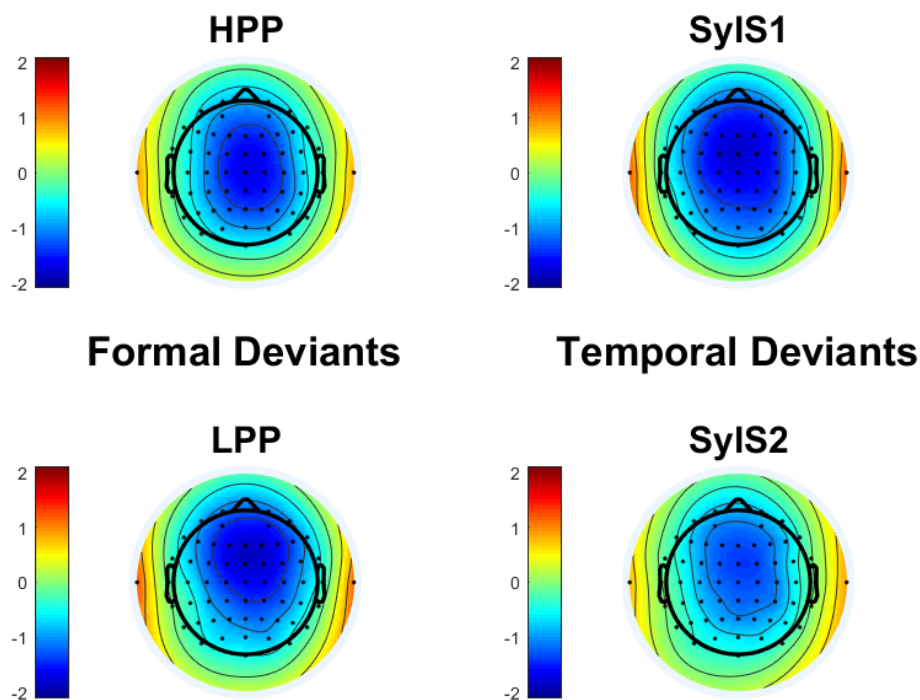
Figure 4: MMN topographies. Average MMN topographies of mean amplitude (+/- 24ms) surrounding individual MMN peaks. Left: Formal deviants averaged across syllable stress. Right: Temporal deviants, averaged across phonotactic probability.

Latency observations were tested in a 2x2 repeated-measures ANOVA (PhonProb x SylStr) on MMN peak latency. This analysis revealed a main effect of PhonProb [$F(1,23) = 16.249$, $p_{adj} = 0.0016$], where HPP deviants show a shorter peak latency than LPP deviants (Figure 6a). Thus, we are able to confirm our second hypothesis that the MMN is sensitive to phonotactic probability, however unlike previous studies (Bonte et al., 2005, 2007; Noordenbos et al., 2013) which demonstrated an effect in MMN amplitude, we observe this effect in MMN latency. Neither amplitude nor latency measures show support for our hypothesis that formal deviant processing may be modulated by syllable stress. No other significant effects were observed for amplitude (Supplementary Table S3) or latency (Supplementary Table S4).
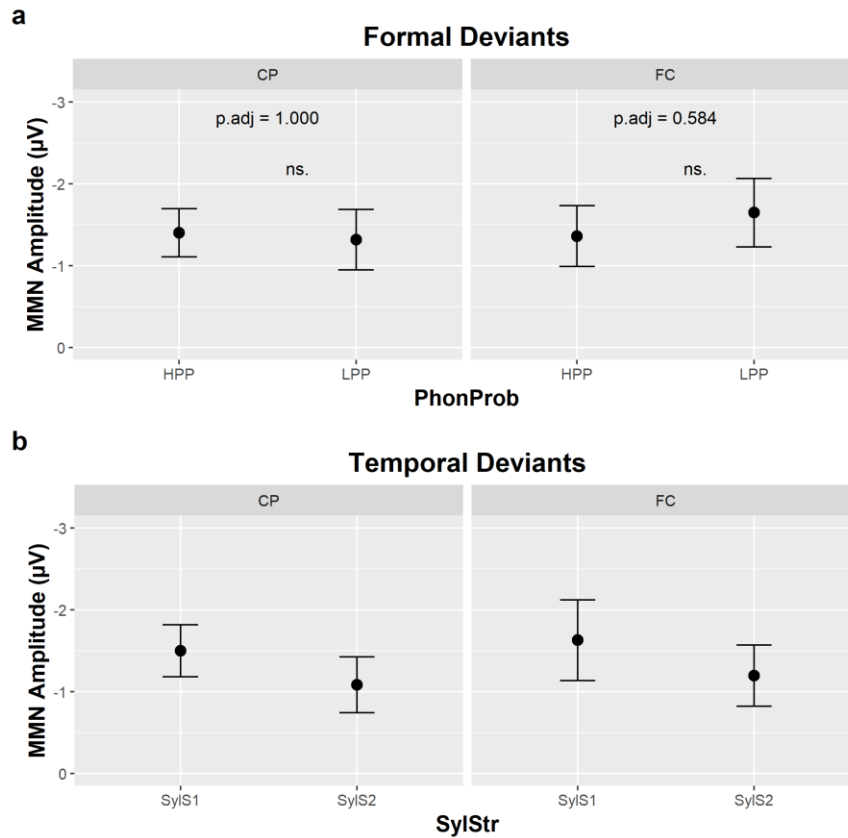
Figure 5: MMN mean amplitudes (Deviant – Standard). (a) Formal deviants, averaged across syllable stress at centroparietal (CP) and frontocentral (FC) ROIs. No significant effect of phonotactic probability on MMN mean amplitude in either ROI (posthoc two-sided paired samples t-test, Bonferroni correction). (b) Temporal deviants, averaged across phonotactic probability at CP and FC. No statistical comparison was made due to non-significant SylStr x Cond interaction. Errorbars represent 95% confidence interval of the mean.
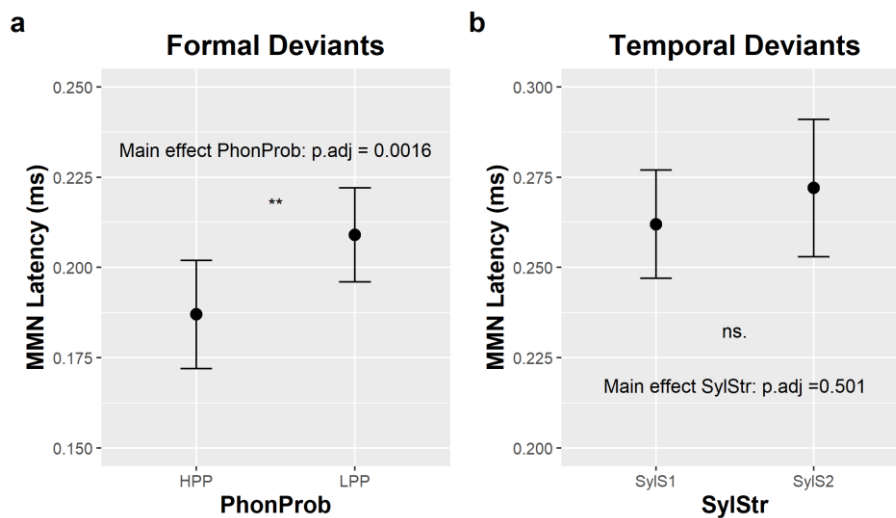
Figure 6: MMN latency measures. (a) Formal deviants averaged across syllable stress. Significant main effect of phonotactic probability (2x2 ANOVA, Bonferroni-Holm correction), with HPP deviants eliciting and earlier MMN compared to LPP deviants. (b) Temporal deviants averaged across phonotactic probability. No significant main effect of syllable stress (2x2 ANOVA, Bonferroni-Holm correction).

3.2   Temporal deviants

Grand average ERPs for standards and temporal deviants time-locked to the onset of the stimulus at the frontocentral ROI are shown in Figure 7. First and second syllable stress temporal deviants, averaged across phonotactic probability, showed a more negative peak response compared to the identical stimulus presented as a standard in a window around 200 – 300 ms after stimulus onset, with comparable topographies (Figure 4b). Visual inspection of the difference waves suggests an effect of syllable stress, with a larger mismatch response for temporal deviants with first syllable stress (top panel), but no apparent effect of phonotactic probability (bottom panel).

These observations were again tested statistically with a 2x2x2x2 repeated-measures ANOVA (PhonProb x SylStr x ROI x Cond) on MMN mean amplitudes. The analysis of mean amplitudes revealed a significant main effect of Cond [$F(1,23)$ = 78.159, p_{adj} < 0.001], confirming our hypothesis that the temporal deviants elicit an MMN. The SylStr x Cond interaction did not reach significance after Bonferroni-Holm correction [$F(1,23)$ = 5.360, p_{adj} = 0.420] (Figure 5b). No other amplitude measures were significant (Supplementary Table S7). The 2x2 ANOVA (PhonProb x SylStr) on MMN peak latency at FCz revealed no significant main effects or interactions (Figure 6b, Supplementary Table S8).
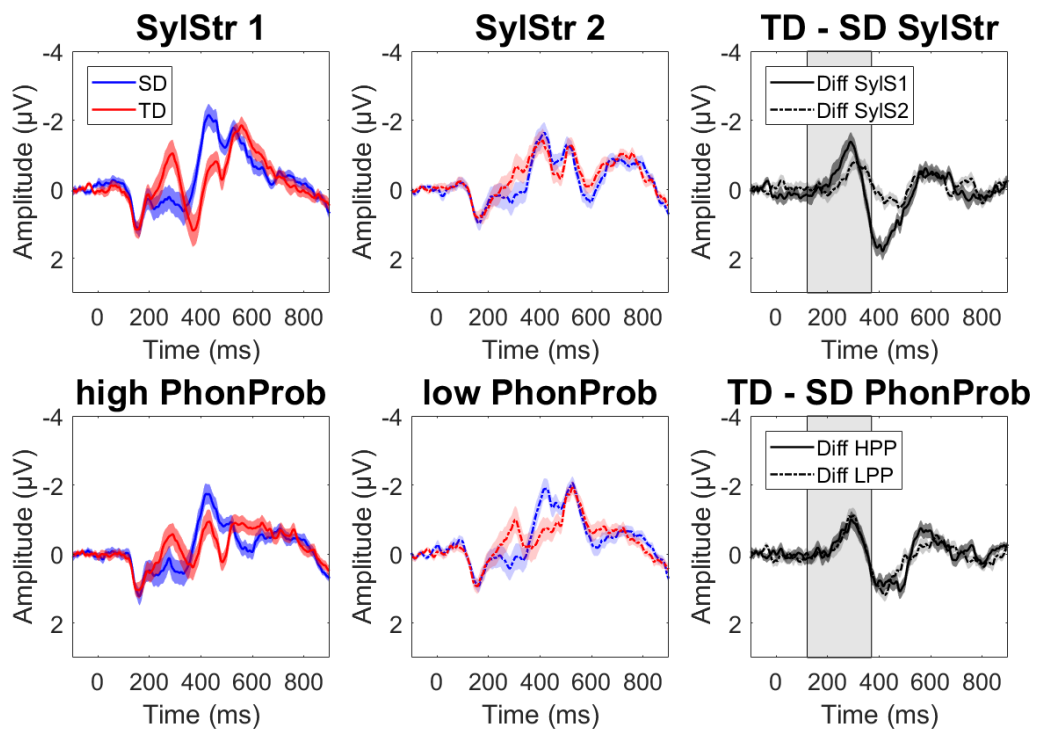
Figure 7: Temporal deviants. Grand average waveforms +/- SEM at frontocentral ROI, time-locked to word onset. Top panel: averaged across phonotactic probability: first syllable stress (red = deviant, blue = standard), second syllable stress, and difference waves (dark grey = SylStr1, light grey = SylStr2). Bottom panel: averaged across syllable stress: high phonotactic probability, low phonotactic probability, and difference waves (dark grey = HPP, light grey = LPP). Shaded area in difference waves represents window for MMN peak extraction.

4   Discussion

The aim of the current experiment was to develop and test an EEG paradigm to provide a measure for formal and temporal predictions in speech perception. This was achieved by means of a passive auditory oddball paradigm, with stimuli consisting of Dutch pseudowords varying in their phonotactic probability (formal prediction) and syllable stress pattern (temporal prediction). The component of interest, the mismatch negativity, is a marker of auditory change detection modulated by experience: It is sensitive to higher-level regularities in the speech signal which are acquired during development, including phonotactic probability (Bonte et al., 2005) and syllable stress (Honbolygó & Csépe, 2013). While both features have been studied in oddball paradigms in isolation, these features vary simultaneously in natural speech. We therefore aimed to examine the effect of manipulating them

simultaneously, to examine whether this would yield similar or different effects on the MMN.

Based on previous ERP experiments manipulating phonotactic probability (Bonte et al., 2005, 2007; Noordenbos et al., 2013), we predicted a larger peak amplitude in the MMN response to formal deviants with high phonotactic probability, compared to their low probability counterparts. Our current results, however, do not show a significant effect of phonotactic probability on MMN peak amplitude. Instead, we observed an effect of phonotactic probability on MMN peak latency, with HPP deviants eliciting an earlier MMN than LPP deviants. A shorter peak latency, indicative of facilitated processing, has been found for other ERP components and paradigms used to investigate the neural correlates of phonotactic probability (Hunter, 2013; Pylkkänen et al., 2002; Stockall et al., 2004). Faster neural processing may thus reflect a possible mechanism underlying the previously reported facilitative behavioural effect of high phonotactic probability on speech processing (Vitevitch & Luce, 1998, 1999, 2005). Therefore, although our current results do not show the same pattern reported in previous MMN studies (Bonte et al., 2005), they may be interpreted in the same context, with the shorter peak latency for MMNs to HPP formal deviants suggesting a facilitated change detection. This effect may reflect Hebbian associative learning, where more frequently co-occurring speech sounds have established more stable auditory cortical memory traces, which can then be accessed more readily (Bonte et al., 2005).

Our study differs from previous passive oddball experiments manipulating phonotactic probability in the Dutch language, which used only first syllable stress in their stimuli. This is the most frequent and natural stress pattern, occurring in around 80% of bisyllabic Dutch words, and close to 95% when including only monomorphemic words (i.e. excluding words with unstressed prefixes; determined by query in CELEX database; Baayen et al., 2001). By including the manipulation of syllable stress in our design, we hoped to examine the interaction between these two factors and see whether manipulating them simultaneously would yield similar or different patterns of MMN responses. While we did not find any significant interaction between these factors, we can observe a trend for a larger latency effect in stimuli with second syllable stress. It is furthermore possible that the simultaneous variation of phonotactic probability and syllable stress within a condition interfered with the

processes underlying the findings reported in previous oddball studies (Bonte et al., 2005, 2007; Noordenbos et al., 2013), which may explain the discrepancy in our findings.

For the temporal domain, we predicted variations in the probability of stress patterns to follow a similar pattern to the phonotactic probability, where stress patterns with high probability (i.e. first syllable stress) would elicit a stronger mismatch response compared to the low probability variation (i.e. second syllable stress). While visual inspection of the ERPs (Figure 6) suggests that this is the case, this comparison did not reach significance after correcting for multiple comparisons (p_{adj} = 0.420, p_{uncorr} = 0.030). Previous studies showing an effect of syllable stress on the MMN were performed primarily in languages where second syllable stress is not a legal construction in bisyllabic words, such as Hungarian (Honbolygó et al., 2004; Honbolygó & Csépe, 2013; Ragó et al., 2014) and Finnish (Ylinen et al., 2009). Here, the 'illegal' second syllable stress deviant is reported to elicit a double MMN response, where the first negative peak is thought to reflect the missing stress on the first syllable, and the second negative peak the response to the unexpected stress on the second syllable (Honbolygó et al., 2004; Honbolygó & Csépe, 2013; Ylinen et al., 2009).

While second syllable stress is rare in Dutch, it is nevertheless a legal construction in bisyllabic words, which suggests that native Dutch speakers would also process deviations in stress patterns in a different manner than speakers of Hungarian or Finnish. A study conducted in German, which also allows variability in syllable stress patterns similar to Dutch, demonstrated that both first and second syllable stress deviants elicited an MMN in adults, while infants only showed a significant MMN in response to first syllable stress (Weber et al., 2004). While the study did not statistically compare the MMN amplitudes for different stress patterns to each other, the results suggest that regularities in syllable stress influence speech perception more heavily early in development, where it is exploited to segment the continuous speech signal into words (Thiessen & Saffran, 2003). Once we have acquired the language successfully, this feature may be less relevant to successful speech perception. Given the observations that some developmental language disorders, such as dyslexia, may be associated with impaired sensitivity to syllable stress (Goswami, 2011; Ladányi, Persici et al., 2020; Lallier et al., 2017), future directions may examine this MMN (in)sensitivity in children and adults with dyslexia.

ERP responses to formal and temporal deviants both suggest that the probability of occurrence within a language has an impact on how we process speech, with more probable constructions being processed more readily. In the case of phonotactic probability, we observe this as more rapid change detection in the form of a shorter MMN peak latency. For syllable stress, we do not observe a statistically significant modulation of the MMN. The current ERP analysis focused specifically on the MMN, however future investigations may explore oscillatory patterns underlying these MMN modulations (see review on oscillatory mechanisms underlying predictions by Arnal & Giraud, 2012). For instance, the observed MMN for temporal deviants may be the result of disrupted beta synchronization to the syllable stress. Additionally, examining gamma band modulations of formal deviant processing may shed light on differential processes underlying short vs. long-term predictions (established based on repeating standards preceding the deviant, and the formal structure of the language acquired during development, respectively).

It is worth noting that the relative frequencies for the "high" and "low" probability items in our manipulations differ for phonotactic probability and syllable stress. While the frequency ratio for HPP vs LPP items is 12.09, that for SylStr1 vs SylStr 2 3.96 (Table 1). It is possible that the difference in frequency between more and less predictable items may influence the degree or morphology of the MMN response, which may in part explain why we observe different MMN patterns in response to formal and temporal deviants. However, it is difficult to directly compare these relative frequencies to each other, as they describe the occurrence of distinct linguistic features. Moreover, while many possible combinations of phoneme clusters exist, syllable stress in bisyllabic words remains binary: either the first, or the second syllable is stressed. Nevertheless, it would be interesting to investigate how the systematic variation of phonotactic probability may affect the MMN response, and how this compares to effects of syllable stress, as previous studies have also been limited to a global categorization of "high" vs. "low" probability (Bonte et al., 2005, 2007; Noordenbos et al., 2013).

The lack of clear interaction between phonotactic probability and syllable stress, and the different response patterns for formal and temporal deviants (latency effect vs. no modulation of MMN) may indicate that these features are processed via independent, parallel mechanisms. This notion is supported by previous evidence

from behavioural and neural findings. Processing of formal and temporal regularities in language appear to develop at different time scales. Infants are already sensitive to the rhythmic properties of their native language in the early days after birth (Nazzi et al., 1998), while sensitivity to its phonological structure, including phonotactics, does not emerge until later in the first year (~6-8 months; Jusczyk et al., 1999; Mattys & Jusczyk, 2001; Maye et al., 2002; Thiessen & Saffran, 2003). In adults, both phonotactic probability and syllable stress have been shown to modulate performance on nonword repetition, with both high phonotactic probability and "typical" stress patterns improving performance (Vitevitch et al., 1997). These features are also associated with distinct neural correlates. For example, the processing of the formal and temporal structure of speech operate via separate oscillatory mechanisms, with neural oscillations in high frequency bands associated with phonological and syntactic encoding, and those in lower frequency bands with tracking the rhythmic structure (Arnal & Giraud, 2012; Giraud & Poeppel, 2012). These processes are further thought to engage distinct functional networks, where the formal structure of the signal is associated with the more classical auditory and speech processing pathways, and the temporal event structure is transmitted via the motor system, including cerebellar and supplementary motor areas (Kotz & Schwartze, 2010, 2016).

The paradigm developed here will further aim at studying the role of formal and temporal predictions in individuals with dyslexia who show reduced sensitivity to neural and behavioural measures of these features. Existing theories on the origin of dyslexia implicate deficits in the processing of formal (Ramus, 2001, 2003; Serniclaes et al., 2004; M. J. Snowling, 2000) or temporal (Goswami, 2011; Thomson & Goswami, 2008) structure of speech, or the combination of the two in the formation of cross-modal representations (Blomert, 2011; Van Atteveldt et al., 2007). Testing our paradigm in children with normal and impaired reading development may help characterize differences in the processing of formal and temporal predictions that are critical to fluent reading skills. In addition to shedding light on the mechanisms and neural correlates of language and reading development, these investigations could be valuable in optimizing the training of pre-reading language skills in kindergarten and/or interventions to facilitate the acquisition of reading skills in children with dyslexia.

Author Contributions: AKE, JMC, MB, and SAK designed the experiment, AKE and JMC prepared the materials, AKE collected and processed the data, AKE, MB, BMJ and SAK interpreted the results, AKE wrote the manuscript, JMC, MB, BMJ and SAK refined the manuscript.

Supplementary Materials

Abbreviations:
FD = Formal Deviant
TD = Temporal Deviant
SD = Standard
MeanAmp = Mean Amplitude
HPP = high phonotactic probability
LPP = low phonotactic probability
SylS1 = first syllable stress
SylS2 = second syllable stress
FC = frontocentral ROI
CP = centroparietal ROI

Supplementary Table S1.
Formal deviants: Descriptive statistics of mean amplitude

PhonProb	SylStr	Cond	ROI	Mean (μV)	SD	N
HPP	**SylS1**	**FD**	**CP**	-2.3060	1.50386	24
			FC	-2.5090	1.68411	24
		SD	**CP**	-.8236	1.55834	24
			FC	-1.0558	1.80899	24
	SylS2	**FD**	**CP**	-1.5318	1.70586	24
			FC	-1.5195	1.96688	24
		SD	**CP**	-.2122	1.35312	24
			FC	-.2484	1.50607	24
LPP	**SylS1**	**FD**	**CP**	-2.0267	2.02642	24
			FC	-2.5083	2.54579	24
		SD	**CP**	-.9220	1.88350	24
			FC	-1.0999	2.20176	24
	SylS2	**FD**	**CP**	-2.4064	1.45576	24
			FC	-2.8504	1.90247	24
		SD	**CP**	-.8729	1.13034	24
			FC	-.9649	1.32014	24

Supplementary Table S2
Formal deviants: Descriptive statistics of peak latency

PhonProb	SylStr	Mean	SD	N
HPP	**SylS1**	.198667	.0536743	24
	SylS2	.174500	.0451750	24
LPP	**SylS1**	.202667	.0495077	24
	SylS2	.216000	.0428384	24

Supplementary Table S3.
*Formal deviants: Results of 2x2x2x2 repeated measures ANOVA on MMN amplitudes
(PhonProb x SylStr x Cond x ROI). P-values adjusted with Bonferroni-Holm correction.*

Effects	F (1,23)	p (adj)
PhonProb	5.667	0.388
SylStr	3.217	0.814
Cond	107.642	< 0.001 ***
ROI	4.838	0.456
PhonProb x SylStr	3.510	0.814
PhonProb x Cond	0.187	1.000
SylStr x Cond	0.465	1.000
PhonProb x SylStr x Cond	1.430	0.814
PhonProb x ROI	3.106	0.814
SylStr x ROI	3.387	1.000
PhonProb x SylStr x ROI	0.487	1.000
Cond x ROI	1.897	1.000
PhonProb x Cond x ROI	8.568	0.112
SylStr x Cond x ROI	0.013	1.000
PhonProb x SylStr x Cond x ROI	0.061	1.000

Supplementary Table S4
*Formal deviants: Results of 2x2 repeated measures ANOVA on MMN peak latency
(PhonProb x SylStr). P-values adjusted with Bonferroni-Holm correction.*

Effects	F(1,23)	p (adj)
PhonProb	16.249	0.0016 **
SylStr	0.190	0.667
PhonProb x SylStr	3.342	0.162

Supplementary Table S5
Temporal deviants: Descriptive statistics of mean amplitude

PhonProb	SylStr	Cond	ROI	Mean	SD	N
HPP	SylS1	TD	CP	-.8900	1.49488	24
			FC	-.6217	2.00712	24
		SD	CP	.6516	1.55185	24
			FC	.9468	1.93654	24
	SylS2	TD	CP	-.5859	1.49901	24
			FC	-.6033	1.65122	24
		SD	CP	.3109	1.27811	24
			FC	.4742	1.66308	24
LPP	SylS1	TD	CP	-1.2490	1.49632	24
			FC	-1.1961	2.09801	24
		SD	CP	.2065	1.45071	24
			FC	.4951	1.79586	24
	SylS2	TD	CP	-.9743	1.58432	24
			FC	-.8575	1.63777	24
		SD	CP	.2976	1.70135	24
			FC	.4545	1.82001	24

Supplementary Table S6
Temporal deviants: Descriptive statistics of peak latency

PhonProb	SylStr	Mean	SD	N
HPP	SylS1	.25967	.046482	24
	SylS2	.25533	.069869	24
LPP	SylS1	.26400	.058562	24
	SylS2	.28850	.057536	24

Supplementary Table S7
Temporal deviants: Results of 2x2x2x2 repeated measures ANOVA on MMN amplitudes (PhonProb x SylStr x Cond x ROI). P-values adjusted with Bonferroni-Holm correction.

Effects	F (1,23)	p
PhonProb	1.719	1.000
SylStr	0.013	1.000
Cond	78.159	<0.001 ***
ROI	3.573	0.923
PhonProb x SylStr	0.841	1.000
PhonProb x Cond	0.981	1.000
SylStr x Cond	5.360	0.420
PhonProb x SylStr x Cond	0.332	1.000
PhonProb x ROI	0.048	1.000
SylStr x ROI	1.811	1.000
PhonProb x SylStr x ROI	1.069	1.000
Cond x ROI	2.000	1.000
PhonProb x Cond x ROI	0.095	1.000
SylStr x Cond x ROI	0.020	1.000
PhonProb x SylStr x Cond x ROI	1.258	1.000

Supplementary Table S8

Results of 2x2 repeated measures ANOVA on MMN peak latency (PhonProb x SylStr). P-values adjusted with Bonferroni-Holm correction.

Effects	F(1,23)	p (adj)
PhonProb	2.039	0.501
SylStr	0.895	0.501
PhonProb x SylStr	1.467	0.501

Chapter 3

ERP and time-frequency correlates of phonological and temporal deviants in dyslexic readers

Based on

ABSTRACT

Developmental dyslexia is a neurodevelopmental disorder characterized by impaired reading development. Dyslexic readers may exhibit disrupted processing of sublexical features of speech, associated with a phonological deficit. Sensitivity to sublexical features such as phonological and temporal regularities in speech is thought to be important for developing accurate phonological representations. In the current experiment, we test Dutch-speaking, adult typical and dyslexic readers in a passive EEG oddball paradigm to investigate whether dyslexic readers exhibit atypical processing of formal (phonological) and temporal regularities of speech. We assessed the mismatch negativity (MMN) as a marker for experience-dependent change detection, as well as the associated phase-reset in the theta frequency range. Here, both formal and temporal deviants elicited a significant MMN. The MMN modulation by phonotactic probability, previously observed in controls, was absent in dyslexic readers. We did not observe any sensitivity of the MMN to variations in syllable stress for typical or dyslexic readers. Time-frequency analyses suggest distinct oscillatory processes underlying change detection for formal and temporal deviants, with formal deviants eliciting increased delta/theta ITC, and temporal deviants a decrease in delta/theta ITC. Taken together, our results suggest atypical sensitivity to phonological but not temporal regularities of speech in Dutch adult dyslexic readers.

1   Introduction

Developmental dyslexia is a reading deficit of neurobiological origin, associated with persistent difficulties in fluent word recognition and impaired spelling abilities. Crucially, the impaired reading development is unexpected based on other cognitive abilities and education opportunities (Lyon et al., 2003). Reports of the prevalence of dyslexia range from 3% to values as high as 20% (Wagner et al., 2020). In addition to reading difficulties, individuals with dyslexia often show disrupted processing of sublexical features of speech. Phonological processing in young children has been shown to predict later reading success (Melby-Lervåg et al., 2012; M. Snowling & Hulme, 1994; Sprugevica & HOien, 2003; Vellutino et al., 2004), and may relate to the formation of mappings between phonemes and their associated graphemes necessary for fluent reading.

To acquire accurate phonological representations, the brain may exploit regularities in the speech signal by means of statistical learning (Peperkamp et al., 2006; Pierrehumbert, 2003). Such sensitivity to regularities in the speech signal is crucial not only for developing phonological skills, but can also guide word learning (Saffran et al., 1996), and is thus an important building block of successful language skill development. It has been proposed that dyslexic individuals have impaired statistical learning abilities (Gabay et al., 2015; Zhang et al., 2021, but see Schmalz et al., 2017) which would render them with reduced sensitivity to regularities in the speech signal, such as phonotactic or transitional probabilities. Phonological processing difficulties observed in dyslexia have also been suggested to arise because of impaired temporal sampling of the speech signal, specifically at the level of delta/theta (~1.5 – 10 Hz) range temporal modulations (Goswami, 2011). This proposal stems from the observation that individuals with dyslexia show perceptual difficulties at the prosodic and syllabic level, which occur roughly at the rate of delta/theta frequencies (Goswami et al., 2013; Leong et al., 2011). Thus, individuals with dyslexia may show impaired phase-locking of delta/theta oscillators in the auditory cortex to rises in the speech amplitude envelope. This may result in overspecification of phonetic information of the speech signal as greater weight is given to high frequency temporal modulations captured by gamma oscillators (Goswami, 2011). Processing of speech prosody and syllable rate are closely linked to

general rhythmic skills, thus this deficit may be related to a general deficit in rhythmic processing (Ladányi, Persici et al., 2020).

Regularities in both the phonological and rhythmic domains of speech are thus both relevant for forming accurate phonological representations, and reduced sensitivity to these regularities may be associated with the commonly observed deficits in dyslexia. Here, we focus specifically on regularities in phonotactic probability and syllable stress, as sensitivity to these features is related to the abovementioned statistical learning and successful encoding of low frequency temporal structure of speech, respectively. Within any given language, certain combinations of speech sounds are more likely to occur than others. This distributional probability is described as phonotactic probability. For example, in Dutch the consonant cluster /ts/ is much more common than /tk/, therefore we can say that words containing /ts/ have a higher phonotactic probability than those containing /tk/. Similarly, any given language can be characterized by regularities in the rhythmic or metric structure of speech. Languages such as Hungarian or Finnish have a fixed stress pattern, where the first syllable of a word is always stressed. In contrast, languages such as Dutch and English allow variations in stress patterns, and changes in lexical stress may even lead to changes in meaning (e.g. present, present; Cutler, 2005; Cutler & Van Donselaar, 2001). Crucially, however, in these languages certain stress patterns may still be more likely to occur than others. For example, Dutch bisyllabic words are more likely to have first syllable stress than second syllable stress. In addition to contributing to language development e.g., in word learning (Jusczyk et al., 1999; Saffran et al., 1996), these regularities continue to shape both speech perception and production throughout the lifespan (e.g., Munson, 2001; Vitevitch et al., 1997; Vitevitch & Luce, 2005).

The neural sensitivity to these regularities can be probed by means of electroencephalography (EEG). The mismatch negativity (MMN) is an event-related potential (ERP) component elicited in the passive oddball paradigm. In this paradigm, the participant is presented with a sequence of stimuli (typically auditory), consisting of a frequently occurring standard stimulus and one or more less frequently occurring deviant stimuli, which differ from the standard in terms of a feature of interest. In the passive paradigm, participants are typically instructed to ignore the stimuli and watch a silent film or read a book. The MMN is characterized as an increased negativity in

response to deviant compared to standard stimuli, typically observed at frontocentral channels in a time-window 100 – 250 ms after auditory deviation (Näätänen et al., 2007). As the MMN is elicited in the absence of overt attention towards a stimulus, it is thought to reflect automatic perceptual change detection, and its amplitude and latency indicate how easily the deviant is detected: deviants that are perceptually more distinguishable elicit an earlier or larger MMN. Crucially, this sensitivity is dependent on experience, as indicated by MMN modulations with varying levels of language experience.

A previous experiment investigated the sensitivity of the MMN to regularities in phonotactic probability and syllable stress in typically reading adults (Emmendorfer et al., 2020). Stimuli consisted of four bisyllabic Dutch pseudowords which varied in phonotactic probability (notsal vs notkal) and syllable stress (first vs second syllable stress). In four separate conditions, each stimulus was presented as a standard, as a deviant differing from standard in terms of phonotactic probability (formal deviant), and as deviant different from the standard in terms of syllable stress (temporal deviant). This allowed comparison of identical stimuli across conditions, to examine effects of the features of interest independent of acoustic differences. Both formal and temporal deviants elicited a significant MMN. In the case of formal deviants, MMN latency was modulated by phonotactic probability, with high probability deviants eliciting an earlier MMN compared to low probability deviants, in line with previous studies demonstrating facilitated change detection for high phonotactic probability items in typical readers (Bonte et al., 2005, 2007; Noordenbos et al., 2013). In the case of temporal deviants, neither amplitude nor latency were modulated. This finding contrasts with previous investigations which have demonstrated a sensitivity of the MMN to variations in syllable stress patterns (Honbolygó & Csépe, 2013). However, these studies were typically conducted in languages with fixed stress patterns, and MMN modulation patterns suggested a violation response for illegal stress patterns, as has also been observed for phonotactic violations (Steinberg et al., 2011).

While ERP effects such as the MMN provide a useful and well-studied indicator of auditory change detection, they only provide a limited picture of the underlying processes. Thus, extending the analyses beyond the time domain can provide a more comprehensive overview of (a)typical processing of phonotactic and temporal

regularities in speech. In the time-frequency domain, two measures often analyzed are inter-trial (phase) coherence (ITC), as a measure of phase alignment across trials, and event-related spectral perturbation (ERSP), as a measure of power modulations that are time-locked to critical events. These measures can vary independently from each other. Time-frequency modulations in the oddball paradigm have been gaining attention in recent years, however, the diversity in experimental designs, subject populations, analytical approaches, and outcome measures has led to a range of reported time-frequency modulations across frequency, phase, and power dimensions. The oscillatory correlate underlying the MMN is thought to be a phase reset resulting in increased inter-trial coherence (ITC) in the theta band (Fuentemilla et al., 2008; Hsiao et al., 2009), sometimes also associated with a concurrent increase in theta power. As with the MMN, these theta modulations have been shown to be associated with perceptual discrimination abilities (Bishop et al., 2011; Jin et al., 2014), speech intelligibility (Koerner et al., 2017), and differences in language abilities (Cantiani et al., 2019; Halliday et al., 2014). As these prior studies have been conducted on simpler stimuli such as tones, vowels or single syllables, we cannot currently form precise hypotheses about modulations of the time-frequency measures by the current stimulus features (phonotactic probability and syllable stress patterns). Due to this, we focus the current time-frequency analyses on the main contrast of deviant vs. standard, and do not analyze time-frequency modulations related to variations in stimulus features.

In the current experiment, we apply a passive auditory oddball paradigm to examine the sensitivity of MMN and oscillatory cortical responses to regularities in phonotactic probability and syllable stress in adult dyslexic readers. By comparing physically identical stimuli presented as standard and deviant, we can investigate auditory change detection beyond mere acoustic effects (Bonte et al., 2005; Eulitz & Lahiri, 2004). In addition to extending upon previous findings (Emmendorfer et al., 2020) by applying the current paradigm to a dyslexic population, the present analysis also goes beyond investigating time-domain modulations such as the MMN ERP response, to explore oscillatory modulations associated with phonotactic and metric change detection. Studies investigating time-frequency correlates of the MMN are commonly limited to tones or single syllable stimuli, resulting in theta band modulations (e.g., Fuentemilla et al., 2008; Hsiao et al., 2009). The current design,

employing bisyllabic pseudowords manipulating lexical stress patterns, may also involve modulations in the delta frequency band, which is associated with tracking prosodic features of speech (Rimmele et al., 2021). With the present approach, we aim to investigate whether adults with dyslexia show atypical processing of phonotactic probability and syllable stress compared to typically reading controls by means of a passive oddball paradigm where these features are manipulated simultaneously in Dutch pseudowords. Specifically, we expect to observe a sensitivity of the MMN to phonotactic regularities of speech stimuli, but not rhythmic regularities in typical readers (Emmendorfer et al., 2020) with more typical/predictable structures (high phonotactic probability) being processed more easily, indexed by a shorter peak latency. We expect this sensitivity to phonotactic probability to be reduced, or atypical, in individuals with dyslexia (Bonte et al., 2007; Noordenbos et al., 2013). We also explore whether Dutch adult dyslexic readers show atypical sensitivity to syllable stress regularities. Finally, we investigate the time-frequency correlates of auditory change detection in relation to these features, where prior research has observed increased theta ITC for deviants relative to standards (Fuentemilla et al., 2008; Hsiao et al., 2009).

2   Methods

2.1   Participants

17 dyslexic readers (mean age = 21.6, standard deviation = 3.54, 14 female), recruited within the Maastricht University community, gave their informed consent to participate in the experiment. Participants were included based on self-reported prior diagnosis of dyslexia, with mean age of diagnosis 10.9 years (standard deviation = 2.94). 17 typically reading controls (mean age = 22.8, standard deviation = 3.80, 11 female) were sourced from a previously published dataset employing the same paradigm (Emmendorfer et al., 2020). The control participants consisted of a random subset of participants who had completed the previous EEG study (n = 24) and for whom behavioral data measures described below were available (n = 21; recorded RAN and EMT performance missing in 3 participants due to technical errors in recording). The final samples included in the current analyses did not differ significantly from each other in age (t(31.8) = 0.982, p = 0.334). The study was approved by the Ethics Committee of the Faculty of Psychology and Neuroscience at

Maastricht University (ERCPN-OZL 205_17_03_2019) and performed in accordance with the approved guidelines and the Declaration of Helsinki.

2.2   Data acquisition

2.2.1 Behavioral tests

Participants completed rapid automatized naming (RAN; Van den Bos & Lutje Spelberg, 2010) and a one-minute word reading test (een-minuut test, EMT; Brus & Voeten, 1973) to assess their phonological and reading skills, as well as a set of subtests from the BAASTA test battery (Dalla Bella et al., 2017) for assessing rhythmic perception and production. The RAN test consisted of 4 subtests (colors, numbers, pictures, letters). The EMT test consists of a list of 120 Dutch words ordered in increasing phonological complexity. Here, participants were instructed to read the words as fast and as accurately as possible before the time limit (1 min) was up. The responses for both RAN and EMT were recorded via a microphone for offline evaluation. For the RAN, the current analysis focusses on the letters and numbers subtests, as these were recently shown to correlate with neural markers of statistical learning in Dutch adult dyslexic and typical readers (Zhang et al., 2021) For the BAASTA, the current analysis focusses on the subtests of duration discrimination and paced tapping with 450 ms ITI, as dyslexic readers often show atypical performance in these tasks (Ladányi, Persici et al., 2020; Thomson et al., 2006; Thomson & Goswami, 2008; Wang et al., 2012)

2.2.2 Oddball paradigm

The current experiment employed a passive auditory oddball paradigm. Stimuli consisted of four Dutch pseudowords (Figure 1a) varying in phonotactic probability (high phonotactic probability = notsal, low phonotactic probability notkal; adapted from Bonte et al., 2005) and syllable stress (first vs second syllable stress), spoken by a female native Dutch speaker and recorded in a sound attenuated chamber using GoldWave Digital Audio Editor (sampling rate 44100 Hz, 16 bit; GoldWave Inc., St. John's, NL Canada). A cross-splicing method was used where the syllables of interest were produced in the context of existing bisyllabic Dutch words, to circumvent pronunciation difficulties due to the scarcity of second syllable stress bisyllabic words

64

with a CVCCVC syllable structure in Dutch (see Emmendorfer et al., 2020 for further detail). In four separate conditions (Figure 1b), each stimulus was presented as standard, formal deviant (differing from standard in terms of phonotactic probability) or temporal deviant (differing from standard in terms of syllable stress). Thus, two types of deviants were presented within a given block. This design allows the comparison of identical stimuli across conditions to examine the effects of phonotactic probability and syllable stress on the EEG signal beyond pure acoustic variations (Figure 1c). Each condition contained a total of 1,620 trials, with 1,332 standard trials, 144 formal deviants, and 144 temporal deviants (8.9% per deviant type). The experiment took place over two sessions. Each session consisted of three runs of approx. 24 mins, split into four blocks (one per condition) of 270 trials. Participants were encouraged to take breaks as needed in between blocks and runs. Within a block, trials were presented with trial duration of 1,200 ms (i.e., inter-stimulus-interval 600 ms). The stimuli were presented in pseudorandom order, with deviants separated by 1-8 standards.
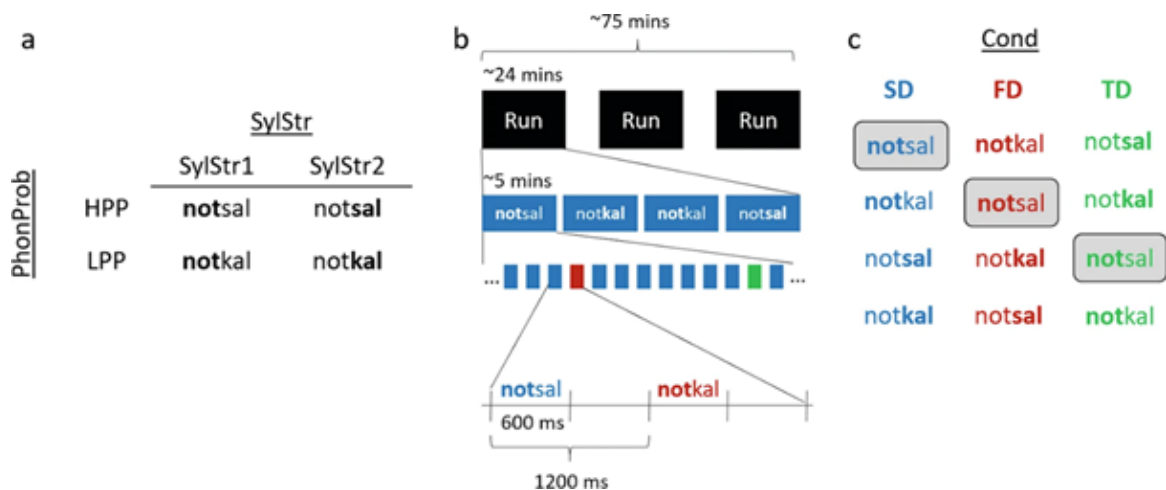
Figure 1: Experimental design as in Chapter 2. (a) Pseudoword stimuli varying in phonotactic probability (PhonProb) and syllable stress (SylStr). Bold font indicates stressed syllable (SylStr1 = first syllable, SylStr2 = second syllable. The phoneme combination '-ts-' constitutes high phonotactic probability (HPP) and '-tk-' low phonotactic probability (LPP). (b) Overview of experimental session. (c) Each stimulus is presented as standard (SD), formal deviant (FD), and temporal deviant (TD) in separate conditions (Cond), allowing the comparison of identical stimuli across conditions (example highlighted for **not**sal).

2.2.3 EEG data acquisition

EEG was recorded with BrainVision Recorder (Brain Products, Munich, Germany) using a 63-channel recording setup. Ag/AgCl sintered electrodes were mounted in an EasyCap electrode cap (EASYCAP GmbH, Herrsching, Germany) according to the 10% equidistant system, including 57 scalp electrodes, left and right mastoids for offline re-referencing, and four EOG electrodes to facilitate removal of artefacts caused by eye movements. The skin at electrode sites was prepared with NuPrep Skin Prep Gel (DO Weaver and Co., USA) and an electrolyte gel was used to keep impedances below 10kΩ. Data were recorded at a sampling rate of 1000Hz, using Fpz as an online reference and AFz as ground. During recording, participants were seated on a comfortable chair in an acoustically and electrically shielded booth and instructed to watch a silent nature documentary while ignoring the auditory stimuli. Participants were encouraged to take breaks in between blocks and runs, to remain awake and relaxed throughout the experiment.

2.3 Analysis

2.3.1 EEG data processing

For ERP analyses, EEG was processed following the procedure used on the control dataset (Emmendorfer et al., 2020). Preprocessing was performed using MATLAB 2017a and the EEG analysis toolbox Letswave 6 (https://github.com/NOCIONS/letswave6). Data were first filtered (band pass 0.5 – 70 Hz, notch filter 48 – 52 Hz) and down-sampled to 250 Hz. Noise from eye-movements, muscle artifacts, and noisy electrodes was removed using independent component analysis (ICA; Delorme et al., 2007) with the runica algorithm implemented in Letswave 6, decomposing the signal into 63 components. The time course recorded during the breaks between blocks was removed from the data to reduce noise prior to the ICA. Artefactual components were selected for removal based on the time course and topography. From the resulting data, -100 to 1000 ms epochs relative to the onset of the stimulus were extracted. After DC removal and baseline correction to the pre-stimulus interval (-100 to 0 ms), an automatic artifact rejection algorithm was applied with an amplitude criterion of 75μV over scalp electrodes to remove trials with remaining high amplitude artifacts, and the data was re-referenced to the average mastoids. Deviants occurring after only one standard

were excluded from analysis. Standards immediately preceding deviants were selected for analysis, resulting in up to 126 trials for each deviant type, and 252 standards for each stimulus. To allow comparing the same standard trials to both formal and temporal deviants, while ensuring equal number of trials across conditions, a random subset of standards was selected to equal the smallest number of deviants across conditions per participant. Within participants, the number of trials was equalized across conditions, leading to a final number of 99 – 124 trials per condition per participant.

An adapted processing pipeline was used for the time-frequency analysis. Here EEG data were processed using the EEGLAB toolbox (version 2019.1, Delorme & Makeig, 2004) in the Matlab environment (version 2019a). Data were filtered 0.5 – 48 Hz and downsampled to a sampling rate of 250 Hz. Noisy channels were identified with the clean_rawdata plugin, and subsequently removed and interpolated from surrounding channels. The data were next epoched to -0.2 – 1.0 s surrounding stimulus. This epoching includes the full trial duration and thus does not remove any data from within a block, but serves the purpose of removing noisy break intervals between blocks prior to ICA. After ICA decomposition, 2-4 components reflecting eye movements were identified by visual inspection of topography, time course and power spectrum and removed from the data. As muscle artifacts are typically associated with high frequency noise outside of our frequencies of interest, these were not removed for this analysis. The epoched data were then concatenated to allow segmentation of longer epochs for the time-frequency analyses. For subsequent processing steps and analyses, only deviant trials (excluding those occurring after only 1 standard) and standard trials immediately preceding deviants were included. The data of these deviant and standard trials were epoched -4.2 s to 5.0 s with respect to stimulus onset. These longer epochs were used to prevent edge artifacts from the wavelet transform from distorting the windows of interest. Time-frequency decomposition was performed at 30 logarithmically spaced frequencies between 1 and 48 Hz, using a wavelet transform that increased from 3 cycles at the lowest frequencies to 10 cycles at the highest frequencies. Baseline power normalization was performed relative to -0.4 s to -0.2 s prior to stimulus onset to avoid temporal "leakage" of post-stimulus activity into the baseline window (Cohen, 2014).

2.3.2 Statistical analysis

For the ERP analysis, trials were averaged after time-locking to the onset of the auditory deviation, corresponding to stimulus onset for temporal deviants, and the /t/-onset for formal deviants. Difference waves were calculated (deviant – standard of identical stimuli, where the standard was always time-locked to the same moment as the respective deviant). Following the procedure used for the control dataset (Emmendorfer et al., 2020), amplitude measures of the MMN were determined around the FCz peak latency for all other electrodes to ensure the comparison of the same underlying process across the scalp (Luck, 2014). First an automatic algorithm in the Letswave toolbox was used to find a negative peak at FCz in the pre-specified time window (80 - 300 ms after /t/-onset for formal deviants, 120 - 370 ms after stimulus onset for temporal deviants; selected based on Emmendorfer et al., 2020). The waveforms and topography were then visually inspected to confirm the selection, or to adjust it to a more fitting peak that reflected the typical frontocentral MMN topography. Mean amplitudes (+/- 24ms surrounding the peak) were calculated for standard and deviant waveforms at all other electrodes at this latency. We focus our analyses on the frontocentral ROI (Fz, F1, F2, F3, F4, FCz, FC1, FC2, FC3, FC4) that was used in the control dataset (Emmendorfer et al., 2020), based on the frontocentral topography of the elicited mismatch response and comparisons with previous literature (Bonte et al., 2005; Honbolygó & Csépe, 2013; Ylinen et al., 2009).

Statistical analyses of the MMN results were completed in R (version 4.1.0). With functions from the rstatix package (Kassambara, 2019), data were checked for outliers using box-plot methods, normality using Shapiro-Wilks test, homogeneity of variances using Levene's test, and homogeneity of covariances using Box's M-test. The outcomes of these tests are reported in the supplementary materials along with the corresponding descriptive statistics. Where violations of normality assumption were found, these were addressed by transforming the data using the transformTukey() function of the rcompanion package (Mangiafico, 2021), which determines a suitable transformation to maximize Shapiro-Wilk's W statistic. Subsequent tests were performed on transformed and untransformed data, and when these results were in alignment with each other, results from untransformed data were reported to facilitate interpretability of the findings. In the case of extreme outliers present in the data, robust ANOVAs from the WRS2 package (Mair & Wilcox, 2020) were performed

68

to validate the outcome and are reported in the supplementary materials. We note that homogeneity of variances was violated for the MMN mean amplitude for low phonotactic probability first syllable stress items (**not**kal) in both formal and temporal deviants, however ANOVA is robust to violations of this assumption given equal sample sizes between groups (Blanca et al., 2018).

First, to establish the presence of an MMN effect, and whether its magnitude differs across dyslexic and typical readers, we performed 2x2 split-plot ANOVAs on mean amplitudes (+/- 24ms surrounding individual peak) with within-subjects factor Condition (Deviant vs Standard) and between-subjects factor Group (Dyslexics vs Controls), for formal and temporal deviants separately. We next tested whether the MMN mean amplitude (calculated as deviant - standard amplitude) was modulated by phonotactic probability or syllable stress, and whether this modulation differs across groups. This was tested via 2x2x2 split-plot ANOVAs including within-subjects factors PhonProb (high vs. low) and SylStr (first vs. second), and Group (Dyslexic vs. Controls) as a between-subjects factor for formal and temporal deviants separately. By separating the test for the presence of the MMN effect from the test of its modulation, we reduce the number of factors in the ANOVAs which reduces the overall number of tests performed and facilitates interpretation of the results (Luck & Gaspelin, 2017). Finally, we tested whether the MMN latency was modulated by phonotactic probability and syllable stress, and whether this differed across groups. This was tested once again via a 2x2x2 split-plot ANOVAs including within-subjects factors PhonProb (high vs. low) and SylStr (first vs. second), and Group (Dyslexic vs. Controls) as a between-subjects factor for formal and temporal deviants separately. To control FWER, p-values within each ANOVA were adjusted using Bonferroni-Holm correction (Cramer et al., 2016).

For the time-frequency analyses, we focus only on the main contrast of standard vs. deviant. Significant time-frequency differences between standard and formal or temporal deviants were assessed using cluster-corrected permutation analyses following the procedure outlined by Cohen (2014). Permutation analyses were performed in the time-frequency space at a frontocentral ROI (Fz, F1, F2, F3, F4, FCz, FC1, FC2, FC3, FC4) averaged across PhonProb and SylStr and collapsed across groups. For formal and temporal deviants separately, two-sided z-tests were performed on 2000 permutations of condition labels (standard vs. formal or temporal

deviants) across 97 time-points (-400 ms – 1008 ms relative to stimulus onset) and 30 frequencies (1 – 48 Hz, logarithmically spaced). Cluster sizes were determined with Matlab's image processing toolbox, and the threshold size was determined based on the 95th percentile of cluster sizes under the null distribution. For the ITC data this yielded a cluster threshold of 207 time-frequency points for formal deviants, and 189 time-frequency points for temporal deviants, for the ERSP data 151 points for formal deviants and 164 points for temporal deviants.

3　Results

3.1　ERP results

To assess the presence of a mismatch negativity across groups, 2x2 ANOVAs with within-subjects factor Condition (Standard vs. Deviant) and between-subjects factor Group (Dyslexic vs. Control) were conducted for formal and temporal deviants separately. For formal deviants, we observed a significant effect of Condition ($F(1,32)$ = 183.391, p.adj < 0.001, η^2_p = 0.851), with deviants exhibiting a greater negativity compared to standards (Figure 2; Supplementary Tables S1, S2). For temporal deviants, we also observed a significant effect of Condition ($F(1,32)$ = 168.011, p.adj < 0.001, η^2_p = 0.840), again with a greater negativity for deviants compared to standards (Figure 3; Supplementary Tables S3, S4). No main effects of Group or interactions between Group and Condition were significant. These results highlight that deviations in phonotactic probability and syllable stress both elicit an MMN, and the magnitude of this mismatch response did not differ between dyslexics and controls.
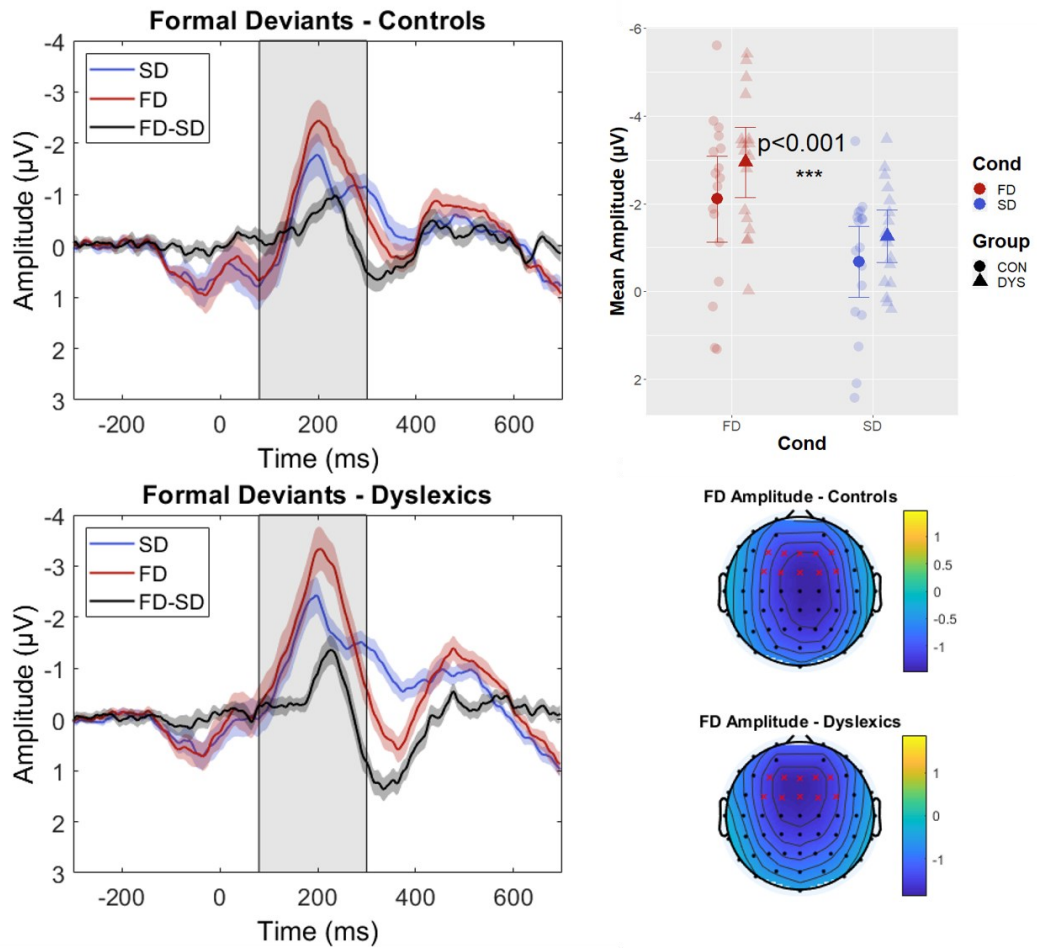
Figure 2: Formal deviants elicited an MMN over frontocentral electrodes (Fz, F1, F2, F3, F4, FCz, FC1, FC2, FC3, FC4). Left: ERP waveforms time-locked to t-onset (auditory deviation) for controls (top) and dyslexics (bottom) at a frontocentral ROI. Right: MMN mean amplitude at +/-24ms surrounding individual peaks (top) and MMN topography (bottom), with ROI channels highlighted in red. FD = formal deviant, SD = standard, CON = Controls, DYS = Dyslexic readers.
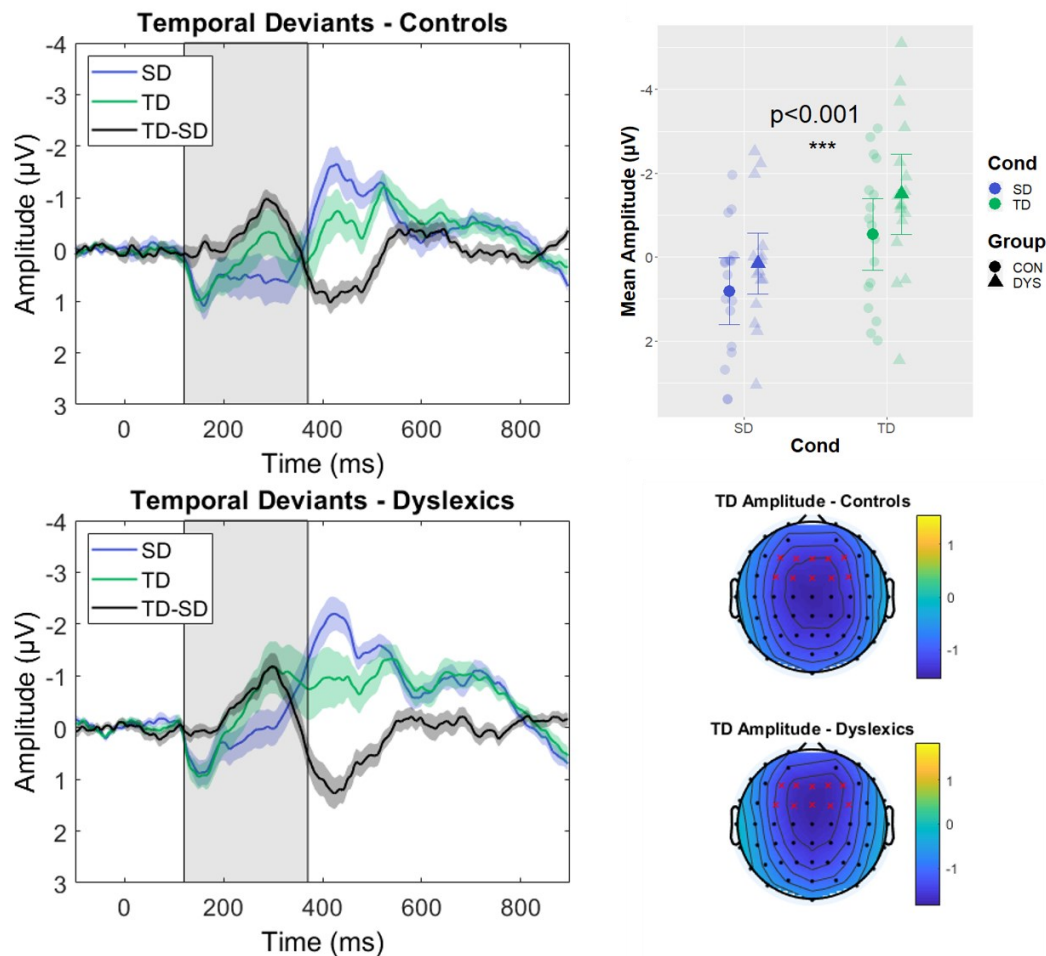
Figure 3: Temporal deviants elicit an MMN over frontocentral electrodes (Fz, F1, F2, F3, F4, FCz, FC1, FC2, FC3, FC4). Left: ERP waveforms time-locked to stimulus onset (auditory deviation) for controls (top) and dyslexics (bottom) at a frontocentral ROI. Right: MMN mean amplitude at +/-24ms surrounding individual peaks (top) and MMN topography (bottom), with ROI channels highlighted in red. TD = temporal deviant, SD = standard, CON = Controls, DYS = Dyslexic readers.

3.1.2 Formal deviants

To test for MMN modulations by the sublexical features of interest and whether they vary across groups, 2x2x2 ANOVAs with within-subjects factors Phonotactic Probability (high vs. low) and Syllable Stress (first vs. second) and between-subjects factor Group (Dyslexics vs. Controls) were conducted on MMN peak latency (determined at FCz) and MMN mean amplitude (calculated as FD - SD of +/-24 ms surrounding individual peak). As MMN peak latency violated the assumption of normally distributed residuals, this ANOVA was performed on transformed and untransformed latency values. As the data transformation did not change the results, we here report the output from untransformed data to facilitate interpretation. No significant main effects or interactions were found (Supplementary Tables S5, S6). Given our previous observation of a significant latency effect in controls, with high

phonotactic probability deviants eliciting an earlier MMN compared to low phonotactic probability deviants (Emmendorfer et al., 2020), we performed post-hoc one-sided paired samples t-tests to test whether this effect was still replicated in the subset of controls included in the current analysis. Indeed, this effect was present in the selection of controls (t(16) = -4.94, p.adj = 0.007, Cohen's d = -0.744) included here, but not in the dyslexic group (t(16) = -0.569, p.adj =0.578, Cohen's d = -0.138). Note that these post-hoc tests were performed on untransformed data as aggregating across syllable stress resolves the violated normality assumption. For MMN mean amplitude, no significant main effects or interactions are reported (Supplementary Table S7, S8; Figure 4).

3.1.3 Temporal deviants

To test for modulations of the MMN by the sublexical features of interest and whether they vary across groups, 2x2x2 ANOVAs with within-subjects factors Phonotactic Probability (high vs. low) and Syllable Stress (first vs. second) and between-subjects factor Group (Dyslexics vs. Controls) were conducted on MMN peak latency (determined at FCz) and MMN mean amplitude (calculated as TD - SD of +/-24 ms surrounding individual peak). Once again, MMN latency violated the assumption of normally distributed residuals, this ANOVA was performed on transformed and untransformed latency values. As the data transformation did not change the results, we here report the output from untransformed data to facilitate interpretation. No significant main effects or interactions were found (Supplementary Tables S9, S10). In the analysis on MMN mean amplitude, we also did not observe any significant main effects or interactions (Supplementary Table S11, S12; Figure 5).
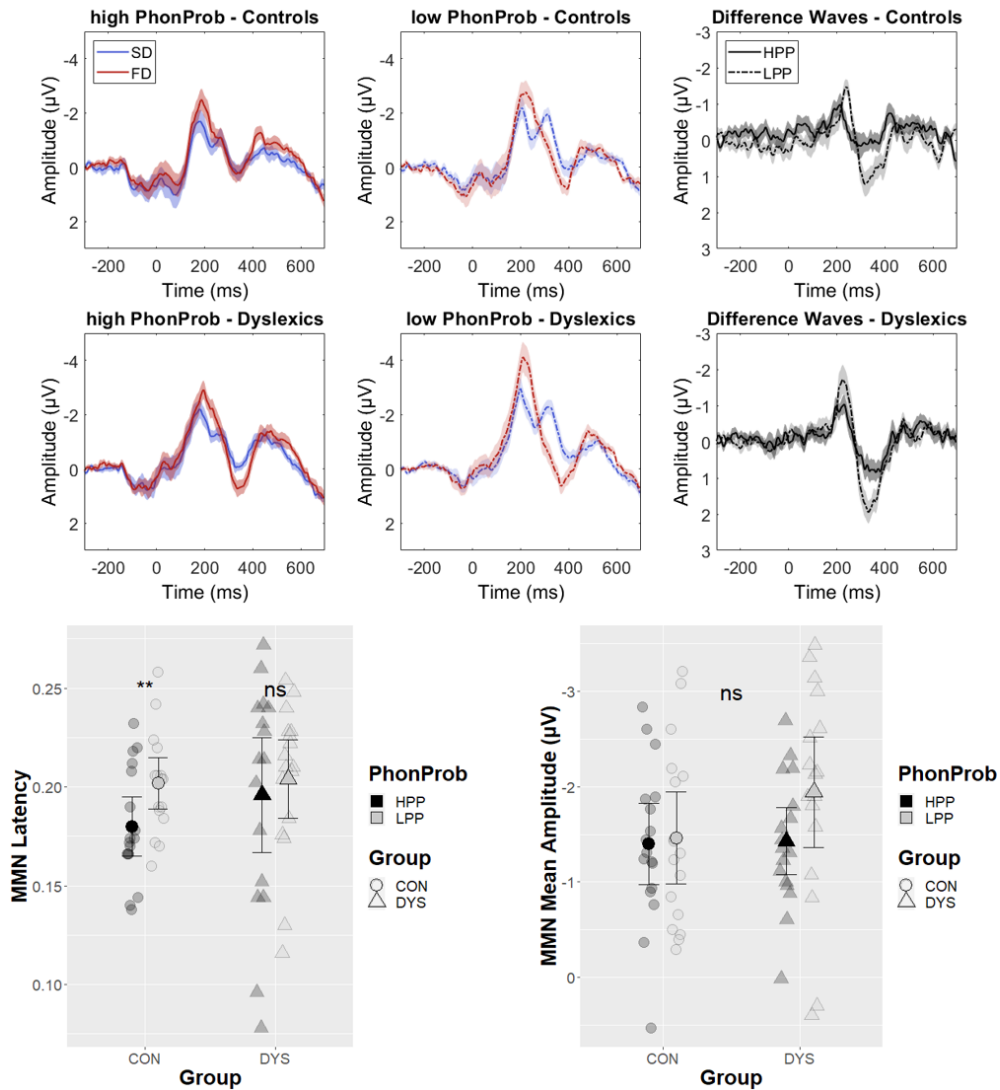
Figure 4: Phonotactic probability modulates MMN latency in typical readers. Top: ERP waveforms time-locked to t-onset (auditory deviation) at a frontocentral ROI (Fz, F1, F2, F3, F4, FCz, FC1, FC2, FC3, FC4), across levels of phonotactic probability and group. Bottom: Effects of phonotactic probability on MMN latency (left) and mean amplitude (right). FD = formal deviant, SD = standard, PhonProb = phonotactic probability, HPP = high phonotactic probability, LPP = low phonotactic probability, MMN = mismatch negativity, CON = Controls, DYS = Dyslexic readers.

74
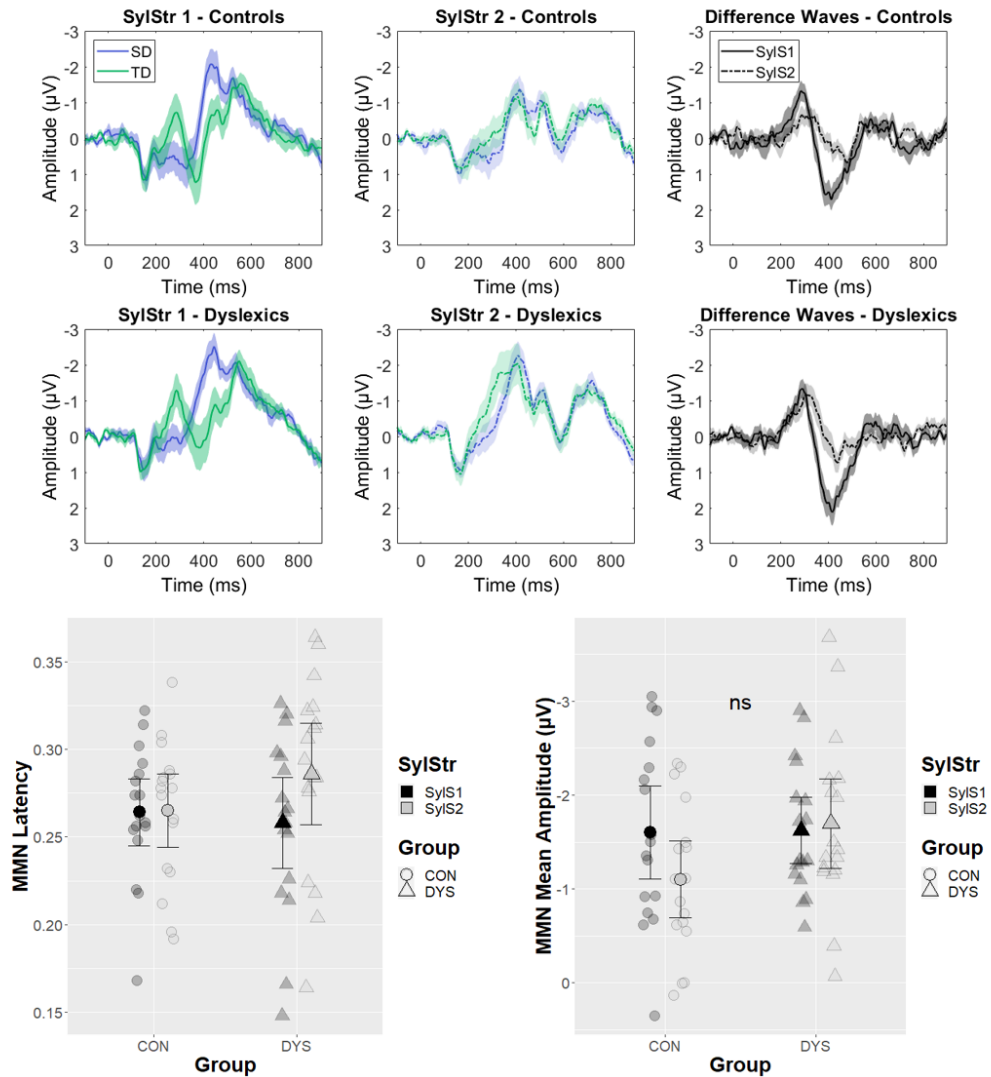
Figure 5: MMN to temporal deviants is not modulated by variations in syllable stress. Top: ERP waveforms time-locked to stimulus onset (auditory deviation) at a frontocentral ROI (Fz, F1, F2, F3, F4, FCz, FC1, FC2, FC3, FC4), across levels of syllable stress and group. Bottom: Effects of phonotactic probability on MMN latency (left) and mean amplitude (right).TD = temporal deviant, SD = standard, SylStr = syllable stress, SylS1 = first syllable stress, SylS2 = second syllable stress, MMN = mismatch negativity, CON = Controls, DYS = Dyslexic readers, ns = non-significant.

3.2   Time-frequency results

To test the time-frequency correlates of auditory change detection in response to formal and temporal deviants, we performed cluster-corrected permutation analyses across the time-frequency spectrogram at a frontocentral ROI, comparing deviants to standards (within-subjects comparison across both groups, averaged over phonotactic probability and syllable stress). This analysis revealed significant ITC modulations for both formal and temporal deviants. Formal deviants elicited an increase in delta/theta ITC and a concurrent increase in ERSP starting at approximately 200 ms after stimulus onset (Figure 6) Temporal deviants were

associated with a decrease in delta/theta ITC starting already around stimulus onset, but no significant ERSP modulations (Figure 7). We next tested whether these modulations differed between controls and dyslexics with two-sided t-tests on the average ITC and ERSP values in the clusters. This analysis revealed a significant difference in ITC between dyslexics and controls for formal deviants (t(32) = -3.10, p = 0.004, Cohen's d = -1.06), but not for ERSP (t(32) = -0.717, p = 0.478, Cohen's d = -0.246). There were no group differences in ITC for temporal deviants (t(32)=-0.762, p =0.451, Cohen's d = -0.262). Visual inspection of the ERSP plots (Figures 6, 7, bottom panels) suggests two distinct time-frequency clusters for typical and dyslexic readers, with typical readers exhibiting power modulations in a lower frequency range (1-3 Hz) than dyslexic readers (2-5 Hz).
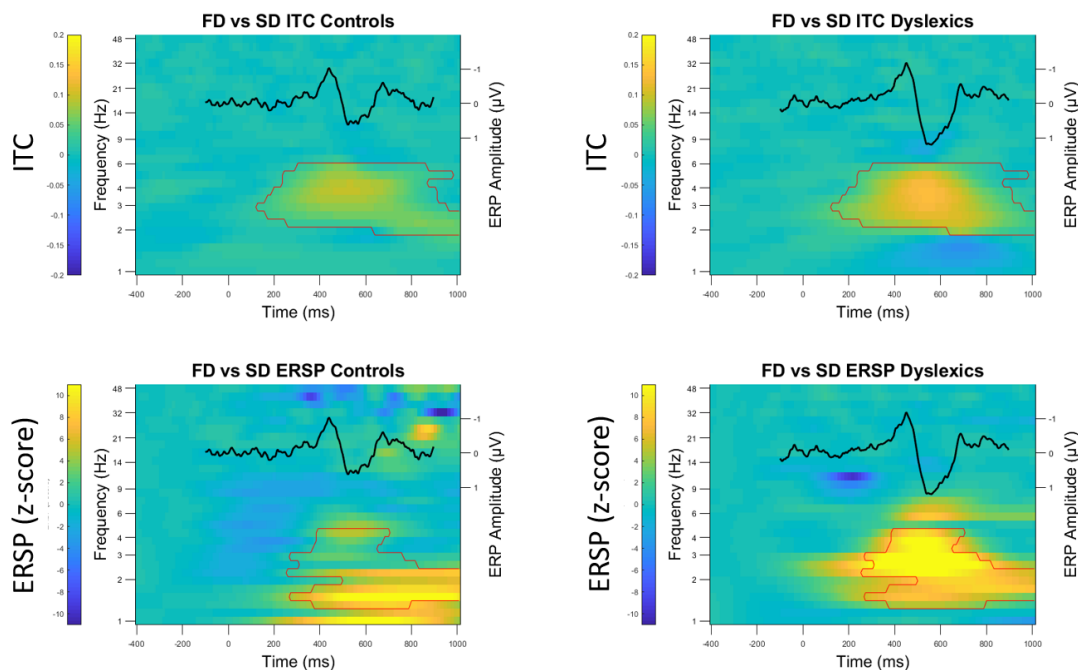
Figure 6: Time frequency correlates of perceptual change detection for formal deviants. Time-frequency plots represent the difference between deviant and standard of identical stimuli averaged across phonotactic probability and syllable stress (left: controls, right: dyslexics) at a frontocentral ROI, averaged across stimuli. In black, the average ERPs (time-locked to stimulus onset in all plots) are overlaid to allow comparison of time-courses. The red contour indicates significant cluster-corrected time-frequency differences between standards and deviants calculated across groups. FD = formal deviants, SD = standards, ITC = inter-trial (phase) coherence, ERSP = event-related spectral perturbation.
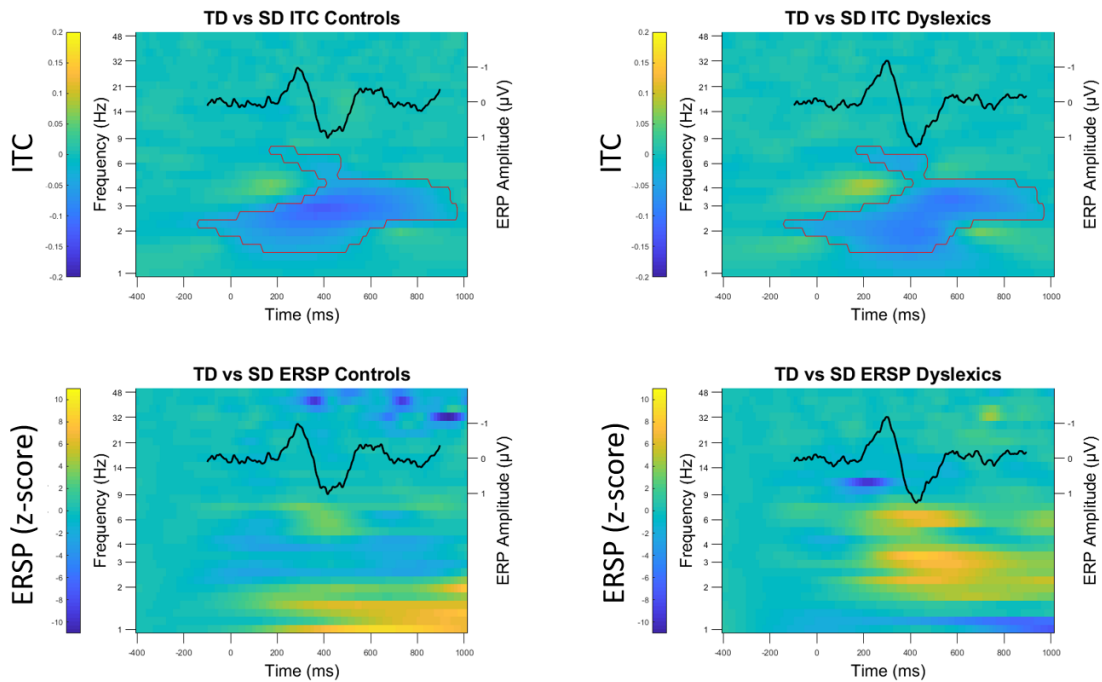
Figure 7: Time frequency correlates of perceptual change detection for temporal deviants. Time-frequency plots represent the difference between deviant and standard of identical stimuli averaged across phonotactic probability and syllable stress (left: controls, right: dyslexics) at a frontocentral ROI, averaged across stimuli. In black, the average ERPs (time-locked to stimulus onset in all plots) are overlaid to allow comparison of time-courses. The red contour indicates significant cluster-corrected time-frequency differences between standards and deviants calculated across groups. TD = formal deviants, SD = standards, ITC = inter-trial (phase) coherence, ERSP = event-related spectral perturbation.

3.3 Behavioral results

Dyslexic readers performed significantly worse than controls in the word reading test ($t(32) = 3.96$, $p_{adj} < 0.001$, Cohen's d = 1.36), but did not differ from controls in their performance on the RAN of letters ($t(32) = -1.47$, $p_{adj} = 0.228$, Cohen's d = -0.503) or numbers ($t(32) = -1.64$, $p_{adj} = 0.167$, Cohen's d = -0.561) , or in paced tapping ($t(30) = 0.502$, $p_{adj} = 1$, Cohen's d = 0.178) or duration discrimination ($t(28) = -0.981$, $p_{adj} = 1$, Cohen's d = -0.358) subtests. As word reading and RAN scores were strongly correlated (Table 1), correlation analyses with EEG responses were only performed with word reading as a behavioral measure of reading skills, as well as the paced tapping and duration discrimination tasks as behavioral markers of rhythmic production and perception. We performed Pearson correlations between number of words read in one minute, paced tapping (coefficient of variability), and duration discrimination (threshold in ms) and EEG measures. Specifically, we correlated the behavioral scores with (1) the overall MMN amplitudes for formal and temporal deviants, as well as with (2) the latency effect in formal deviants, calculated as LPP

latency - HPP latency, and (3) the ITC increase for formal deviants relative to standards. No significant correlations were observed (Table 2).

Table 1

Behavioral results

Task	Descriptive stats Mean (SD)		Group difference	Pearson Correlation (r)			
	Controls	Dyslexics	t (df)	RAN Letters	RAN Number	Tapping	Duration
EMT	96.1 (10.1)	78.8 (14.9)	**3.96 (32) *****	**-0.49 ***	**-0.52 ***	0.06	-0.15
RAN Letters	20.8 s (2.98)	22.6 s (3.97)	-1.47 (32)	1.00	**0.67 *****	-0.10	0.03
RAN Numbers	21.1 s (2.82)	22.9 s (3.82)	-1.64 (32)		1.00	-0.19	-0.03
Tapping n=16/group	0.056 (0.015)	0.053 (0.011)	0.502 (30)			1.00	0.27
Duration n=15/group	28.5 ms (12.7)	33.3 ms (14.1)	-0.981 (28)				1.00

Note: EMT = words read in 60 s, RAN Letters/Numbers = time (s) to name 50 letters/numbers), Tapping = Paced tapping (BAASTA subtest) coefficient of variability for paced tapping to isochronous tones with ITI of 450 ms, n = 16 per group due to missing data, Duration = Duration discrimination (BAASTA subtest) threshold (ms), n = 15 per group due to missing data. *p<.05, *** p<0.001. p-values are Bonferroni-corrected.

Table 2

Correlation of EEG and behavioral results (Pearson's r)

	FD MMN amplitude	TD MMN amplitude	LPP – HPP MMN latency	FD – SD ITC
EMT	0.056	0.088	0.17	-0.22
Tapping	-0.31	-0.28	-0.1	-0.1
Duration	0.11	-0.033	-0.21	-0.053

Note: EMT = words read in 60s, Tapping = Paced tapping (BAASTA subtest) coefficient of variability for paced tapping to isochronous tones with ITI of 450 ms, n = 16 per group due to missing data. Duration = Duration discrimination (BAASTA subtest) threshold, n = 15 per group due to missing data. Values represent Pearson's r. No correlations were significant before Bonferroni correction.

4   Discussion

The current experiment set out to examine whether adults with dyslexia show atypical implicit processing of phonotactic probability and syllable stress using a passive oddball paradigm. Previous research showed that the MMN is sensitive to variations in phonotactic probability (Bonte et al., 2005; Emmendorfer et al., 2020), but that this sensitivity is reduced in both children (Bonte et al., 2007) and adults (Noordenbos et al., 2013) with dyslexia. Additionally, the MMN has been reported to be sensitive to variations in syllable stress in languages with fixed stress patterns such as Hungarian (Honbolygó & Csépe, 2013) and Finnish (Ylinen et al., 2009), but not in languages with variable stress patterns such as Dutch (Emmendorfer et al., 2020). While previous studies have manipulated these features in isolation, we employ a multi-feature approach, investigating the sensitivity of the MMN to phonotactic probability and syllable stress when they are manipulated simultaneously. To further elucidate whether formal and temporal deviance detection employs similar neural substrates, we perform a time-frequency analysis.

Formal deviants, which differed from the standard in terms of phonotactic probability, elicited a significant MMN, indexed by a larger negativity for deviants compared to standards over fronto-central electrode sites. The magnitude of this MMN did not differ between dyslexics and controls. While the initial ANOVA analysis did not reveal any significant MMN modulation by stimulus features (phonotactic probability or syllable stress) or group, post-hoc t-tests revealed that MMN latency was sensitive to variations in phonotactic probability in typical readers, with high

phonotactic probability deviants eliciting an earlier mismatch response compared to low phonotactic probability deviants as was observed in the full control dataset (Emmendorfer et al., 2020). This observation is consistent with the notion of facilitated change detection for more predictable deviants. In contrast, dyslexic participants did not show a sensitivity of MMN latency to variations in phonotactic probability. Albeit in the absence of a significant difference in the direct group comparison, this is in line with previous reports of reduced sensitivity to phonotactic probability in both children (Bonte et al., 2007) and adults (Noordenbos et al., 2013) with dyslexia. However, these studies reported a modulation of MMN amplitude, rather than variations in latency. This discrepancy can likely be explained by differences in the current design, where previous studies manipulated phonotactic probability in isolation, either in single syllables (Noordenbos et al., 2013) or bisyllabic pseudowords (Bonte et al., 2007) with first syllable stress. With our design manipulating both features simultaneously, processes generating the MMN response may be altered, leading to a different modulation by phonotactic probability (Emmendorfer et al., 2020).

Temporal deviants, which differed from the standard in terms of syllable stress, also elicited a significant MMN at frontocentral electrode sites. As with the formal deviants, the magnitude of this MMN response did not differ between dyslexics and controls. We also did not observe any further modulations of this MMN by stimulus features in either participant group. This stands in contrast to previous studies investigating modulations of the MMN to variations in stress patterns (e.g., Honbolygó & Csépe, 2013; Schaadt & Männel, 2019), however there are several noteworthy differences between prior approaches and the current experiment. The current approach applies a reversed MMN design, where the MMN difference waves illustrate the difference between deviant and stimulus of *physically identical stimuli* (Bonte et al., 2005; Eulitz & Lahiri, 2004). This minimizes the influence of acoustic differences between standard and deviant on the MMN. The majority of studies investigating MMN sensitivity to syllable stress patterns compare deviants and standards of different stimuli as they are presented in the blocks (e.g., second syllable stress deviant - first syllable stress standard), thus acoustic stimulus differences likely also contributed to the reported effects. Furthermore, many studies investigating the MMN sensitivity to syllable stress variations have been conducted in languages with

a fixed stress pattern (e.g., Honbolygó et al., 2004; Honbolygó & Csépe, 2013; Ylinen et al., 2009), while those studies conducted in languages with variable stress pattern either focus on MMN sensitivity to specific acoustic markers of stress (e.g., pitch, duration, amplitude; Tong et al., 2014; Zora et al., 2015), or do not directly statistically compare the MMN elicited to different stress patterns, but simply note its presence or absence (Weber et al., 2004), or group differences within a condition (Schaadt & Männel, 2019). One study that did analyze the MMN by comparing physically identical stimuli reported an increased MMN to second syllable stress deviants (Honbolygó & Csépe, 2013). However, as this study was conducted in Hungarian, a language with obligatory first syllable stress, this pattern is more in line with a violation response, as the second syllable stress deviant stimuli here violate the rules of the language. Similar patterns of enhanced MMN amplitudes have been observed for violations of phonotactic constraints (Steinberg et al., 2011), standing in contrast to patterns observed with variations in probability of legal phonotactic structures (Bonte et al., 2005; Emmendorfer et al., 2020; Noordenbos et al., 2013).

An additional aim of the current study was to investigate time-frequency correlates of auditory change detection of formal and temporal deviants. Here, we focused on the main contrast of deviant vs. standard rather than investigate modulations related to phonotactic probability and syllable stress, as the lack of prior studies investigating time-frequency modulations in oddball paradigms with similarly complex stimuli made it difficult to formulate precise hypotheses about expected modulations. The MMN is thought to be generated by a phase-reset in the theta frequency range (Fuentemilla et al., 2008; Hsiao et al., 2009), which is interpreted as a marker for sensory memory processes underlying deviance detection, as the deviant stimulus is compared to the sensory memory trace of the standard. In line with this, the current data show an increase in delta/theta ITC in response to formal deviants compared to standards. In contrast, temporal deviants are associated with a decrease in delta/theta ITC relative to standards. This finding is particularly interesting, as the perceptual markers of stress such as duration (Fuentemilla et al., 2008; Hsiao et al., 2009), pitch (Halliday et al., 2014), and intensity (de la Salle et al., 2019) are generally associated with the typical increase in theta ITC. The current observation points more towards a disruption of the tracking of the regular speech rhythm over the course of a block. These contrasting observations for

formal and temporal deviants in the time-frequency domain highlight that regularities in the phonotactic and rhythmic structure of speech are processed via dissociable cortical mechanisms, which may also explain the lack of interactions observed between phonotactic probability and syllable stress in the ERP domain.

A further interesting observation in the time-frequency analysis is the increased delta/theta ITC to formal deviants, averaged across low and high phonotactic probability deviants, in dyslexic individuals compared to controls. Such a pattern has been previously observed in children with dyslexia in an oddball paradigm using frequency deviants in simple tone stimuli (Halliday et al., 2014). Our findings, together with those by Halliday and colleagues, suggest that individuals with dyslexia need not show impairments in phase-locking in the delta/theta range. This finding does not seem to be in line with the notion of disrupted phase-locking in the delta and theta range as a neurobiological origin of dyslexia, as proposed by Goswami (2011). However, while not disrupted per se, enhanced ITC in the dyslexic group may still be interpreted as *atypical* phase-locking. The increased ITC we observe in dyslexic adults may suggest more effortful processing of the perceptual deviance, as is observed in MMN studies with linguistic violations (Honbolygó & Csépe, 2013; Steinberg et al., 2011). In their younger subgroup of dyslexic children, Halliday and colleagues (2014) only observed this increased theta ITC for small frequency deviants (1030 Hz vs. 1000 Hz) but not large frequency deviants (1200 Hz vs. 1000 Hz), which also supports the notion of more effortful processing. The fact that we do not observe group differences in the MMN amplitude may indicate that time-frequency measures may provide a more sensitive index of the underlying neural processes. Indeed, the MMN amplitude for formal deviants is larger for dyslexics than controls (Supplementary Table S1), but this difference does not reach significance (Supplementary Table S2).

Given the temporal overlap of the increased delta/theta ITC with both the MMN and the P3 window (Figure 6), it is also possible that this effect reflects an involuntary attention switch to evaluate the perceptual change (Fitzgerald & Todd, 2020), which may be increased in the dyslexic group. Indeed, increased theta ITC has been previously associated with an attentional shift in participants hearing their own name (Tamura et al., 2015), and some evidence points toward impaired auditory attentional interference control in dyslexic readers (Gabay et al., 2020), which may

render them more susceptible to involuntary attention switching. This differs from recent reports of decreased attentional shifting to upcoming stimuli in dyslexic children in an audiovisual oddball paradigm (Meyer & Schaadt, 2020). However, this study investigated pre-stimulus alpha ITC, and required children to direct their attention toward the visual speech stimuli which preceded the auditory onset, thus highlighting impairments in endogenous attentional control, rather than involuntary attention switches. Studies of auditory interference control of attention in dyslexia are rare (Gabay et al., 2020), thus this interpretation warrants follow up investigations.

While we did not observe any group differences in ERSP averaged across the significant time-frequency cluster, visual inspection of the time-frequency plots reveals differences in the frequency composition between dyslexic and typical readers (Figure 6, bottom panel). Here it appears that typical readers exhibit power modulations in a lower frequency range (~1-3 Hz) compared to dyslexic readers (~3-6 Hz), while the time-frequency cluster in the ITC appears to be homogeneous across groups (Figure 6, top panels). This difference in the affected frequency bands across groups suggests that the MMN observed may reflect different cortical processes in typical vs. dyslexic readers, which would not be observable in classical ERP analyses. Further investigations are necessary to determine the meaning of these differences in power modulation.

The sensitivity of the MMN to variations in phonotactic probability has been reported to be correlated with reading measures in children (Bonte et al., 2007), but not adults (Noordenbos et al., 2013) with and without dyslexia. In line with this observation, we also did not find a correlation between reading fluency and the latency effect in our sample. This reported difference between adults and children may be related to compensatory strategies in reading skills during adolescence and adulthood. However, it may also emerge as a result of a sampling bias in the adult dyslexic population. In the current dataset (and possibly also in Noordenbos et al., 2013, though this is not explicitly reported), dyslexic participants were recruited from within the university to match the sample of control participants previously tested, and were included based on self-reported prior diagnosis. Given the comparably high level of reading experience in university students, this may result in improved reading abilities relative to dyslexic individuals who do not pursue university education.

We did not observe any group differences in processing temporal deviants here. There may be multiple reasons for this outcome. Increasing evidence suggests that impairments in processing lexical stress may be related to a decreased explicit awareness of stress patterns, rather than impairments in implicit stress perception tested in the passive oddball paradigm (Anastasiou & Protopapas, 2015; Barry et al., 2012; Mundy & Carroll, 2012, 2013). Perceptual sensitivity to atypical stress patterns (here: second syllable stress) likely emerges within the first few years of life (Schaadt & Männel, 2019; Weber et al., 2004). Evidence from Schaadt & Männel (2019) suggests that this sensitivity, measured in a passive oddball paradigm, is reduced in pre-reading children who later develop reading difficulties. This reduced sensitivity to atypical stress patterns may be simply related to a developmental delay, and not a lasting deficit that persists into adulthood. Finally, the dyslexic participants included in this study were recruited from within the university community based on self-reported diagnosis, and are thus likely not representative of the dyslexic population as a whole (Kortteinen et al., 2021). This may also explain the lack of group differences in the tests of rhythmic skills.

To summarize, our results highlight distinct oscillatory correlates underlying auditory change detection for deviants in phonotactic probability and syllable stress, with formal deviants eliciting the typically observed increase in delta/theta ITC associated with sensory memory processes, and temporal deviants showing a disruption of tracking of the repetitive speech rhythm during a block. This observation suggests that processing of the formal and temporal structure of the speech signal may operate via dissociable cortical mechanisms as suggested by prior models of speech processing (Arnal & Giraud, 2012; Giraud & Poeppel, 2012; Kotz & Schwartze, 2010). The current data further suggest that Dutch adults with dyslexia show a reduced sensitivity to variations in phonotactic probability (no MMN latency modulation), as well as atypical processing of formal deviancy overall (increased theta ITC). In line with prior reports of atypical sensitivity to phonotactic probability (Bonte et al., 2007; Noordenbos et al., 2013), and slower statistical learning in dyslexic readers (Zhang et al., 2021), this observation supports the notion that the ability to acquire a sensitivity to phonological regularities in speech may be important to successful reading development. In contrast, no group differences are observed for variations in syllable stress or temporal deviancy overall. These patterns of results

(for both phonotactic probability and syllable stress) do not correlate with behavioral measures of reading ability. These findings are unexpected given theories proposing disrupted processing of the temporal structure of speech, arising from impaired phase-locking to slow amplitude modulations in the speech signal in individuals with dyslexia (Goswami, 2011). A cross-linguistic perspective may resolve the diversity of findings concerning the role of linguistic rhythm in dyslexia (Lallier et al., 2017). According to this framework, speakers of languages with variable stress patterns, such as Dutch, should be particularly sensitive to variations in stress patterns, as it can be an important cue in resolving lexical conflict. While the current data do not support this notion, this might be related to the current task demands, as participants only passively listened to a sequence of pseudowords. Future directions should investigate the developmental trajectories of the sensitivity to these features prior to and during reading acquisition to better characterize their contribution to reading skills.

Supplementary Materials

Supplementary Table 1

Descriptive statistics and ANOVA assumptions for formal deviants MMN.

Cond	Group	Descriptive Stats		Shapiro-Wilk test		Levene's test		Box's M-test	
		Mean (µV)	Std Dev	W	p	W	p	M	p
FD	DYS	-2.94	1.55	0.956	0.553	0.317	0.577	0.768	0.381
	CON	-2.11	1.90	0.947	0.405				
SD	DYS	-1.26	1.16	0.965	0.726	0.894	0.351		
	CON	-0.678	1.58	0.927	0.191				

Supplementary Table 2

2x2 split-plot ANOVA on mean amplitude at MMN latency for formal deviants (Bonferroni-Holm corrected)

Effect	$F(1,32)$	p_{adj}	Effect size η^2_p
Group	1.813	0.376	0.054
Cond	183.391	<0.001***	0.851
Group x Cond	1.227	0.376	0.037

Supplementary Table 3

Descriptive statistics and ANOVA assumptions for temporal deviants MMN.

Cond	Group	Descriptive Stats		Shapiro-Wilk test		Levene's test		Box's M-test	
		Mean (µV)	Std Dev	W	p	W	p	M	p
TD	DYS	-1.50	1.85	0.971	0.838	0.036	0.851	0.105	0.746
	CON	-0.54	1.66	0.948	0.423				
SD	DYS	0.159	1.42	0.922	0.163	0.665	0.421		
	CON	0.812	1.56	0.956	0.561				

Supplementary Table 4

2x2 split-plot ANOVA on mean amplitude at MMN latency for temporal deviants (Bonferroni-Holm corrected)

Effect	Classical ANOVA (rstatix package)			Robust ANOVA (WRS2 package)		
	$F(1,32)$	p_{adj}	Effect size η^2_p	df1, df2	$F(df1,df2)$	p_{adj}
Group	2.176	0.3	0.064	1, 17.01	1.772	0.401
Cond	168.011	<0.001***	0.840	1, 19.99	77.401	<0.001***
Group x Cond	1.750	0.3	0.052	1, 19.99	1.386	0.401

Note: Robust ANOVA was performed to validate results due to extreme outliers in data.

Supplementary Table 5

Descriptive statistics and ANOVA assumptions for MMN latency of formal deviants

PhonProb	SylStr	Group	Descriptive Stats		Shapiro-Wilk test		Levene's test		Box's M-test	
			Mean (s)	Std Dev	W	p	W	p	M	p
HPP	**SylS1**	**DYS**	0.192	0.072	0.923	0.166	1.02	0.320	1.03	0.310
		CON	0.187	0.056	0.990	0.999				
	SylS2	**DYS**	0.200	0.055	0.823	**0.004**	0.062	0.805		
		CON	0.174	0.049	0.949	0.436				
LPP	**SylS1**	**DYS**	0.206	0.061	0.925	0.176	0.028	0.868		
		CON	0.193	0.055	0.921	0.156				
	SylS2	**DYS**	0.202	0.053	0.857	**0.012**	0.003	0.959		
		CON	0.210	0.048	0.843	**0.008**				

Note: Values from untransformed data shown here. Due to the violation of the normality assumption, data were transformed using the transformTukey() function from the rcompanion package, which determines a suitable transformation to maximize Shapiro-Wilk's W statistic collapsed across factor levels. Here, data underwent a transformation of $x^{1.95}$.

Supplementary Table 6

2x2x2 split-plot ANOVA on MMN latency for formal deviants (Bonferroni-Holm corrected)

Effect	Untransformed			Transformed		
	F(1,32)	p_{adj}	Effect size η^2_p	F(1,32)	p_{adj}	Effect size η^2_p
Group	0.749	1.000	0.023	1.299	1.000	0.039
PhonProb	3.292	0.553	0.093	3.227	0.574	0.092
SylStr	0.043	1.000	0.001	0.00001	1.000	0.0000005
Group x PhonProb	0.614	1.000	0.019	0.875	1.000	0.027
Group x SylStr	0.001	1.000	0.00003	0.024	1.000	0.0008
PhonProb x SylStr	0.284	1.000	0.009	0.408	1.000	0.013
Group x PhonProb x SylStr	1.56	1.000	0.046	1.495	1.000	0.045

Note: Data were transformed using the transformTukey() function from the rcompanion package, which determines a suitable transformation to maximize Shapiro-Wilk's W statistic. Here, data underwent a transformation of $x^{1.95}$.

Supplementary Table 7

Descriptive statistics and ANOVA assumptions for MMN mean amplitude of formal deviants.

PhonProb	SylStr	Group	Descriptive Stats		Shapiro-Wilk test		Levene's test		Box's M-test	
			Mean (μV)	Std Dev	W	p	W	p	M	p
HPP	SylS1	DYS	-1.43	0.934	0.956	0.566	0.308	0.583	0.0253	0.874
		CON	-1.57	1.01	0.926	0.186				
	SylS2	DYS	-1.43	1.20	0.973	0.874	0.362	0.551		
		CON	-1.23	1.52	0.975	0.896				
LPP	SylS1	DYS	-1.68	1.47	0.975	0.897	7.26	**0.011**		
		CON	-0.95	0.602	0.970	0.810				
	SylS2	DYS	-2.21	1.29	0.969	0.382	1.20	0.281		
		CON	-1.97	1.58	0.945	0.797				

Supplementary Table 8

2x2x2 split-plot ANOVA on MMN mean amplitude for formal deviants (Bonferroni-Holm corrected)

Effect	F(1,32)	p$_{adj}$	Effect size η²$_p$
Group	1.227	1.000	0.037
PhonProb	1.865	0.996	0.055
SylStr	2.011	0.996	0.059
Group x PhonProb	1.137	1.000	0.034
Group x SylStr	0.028	1.000	0.0009
PhonProb x SylStr	5.952	0.140	0.157
Group x PhonProb x SylStr	1.160	1.000	0.035

Supplementary Table 9

Descriptive statistics and ANOVA assumptions for MMN latency of temporal deviants

PhonProb	SylStr	Group	Descriptive Stats		Shapiro-Wilk test		Levene's test		Box's M-test	
			Mean (s)	Std Dev	W	p	W	p	M	p
HPP	SylS1	DYS	0.236	0.06	0.887	**0.042**	0.364	0.550	0.382	0.536
		CON	0.262	0.049	0.932	0.232				
	SylS2	DYS	0.287	0.074	0.864	**0.018**	0.039	0.845		
		CON	0.248	0.067	0.944	0.373				
LPP	SylS1	DYS	0.279	0.056	0.952	0.488	0.053	0.820		
		CON	0.266	0.058	0.883	**0.035**				
	SylS2	DYS	0.286	0.062	0.855	**0.013**	0.513	0.479		
		CON	0.281	0.067	0.910	0.101				

Note: Values from untransformed data shown here. Due to the violation of the normality assumption, data were transformed using the transformTukey() function from the rcompanion package, which determines a suitable transformation to maximize Shapiro-Wilk's W statistic collapsed across factor levels. Here, data underwent a transformation of $x^{2.425}$.

Supplementary Table 10

2x2x2 split-plot ANOVA on MMN latency for temporal deviants (Bonferroni-Holm corrected)

Effect	Untransformed			Transformed		
	F(1,32)	p_{adj}	Effect size η^2_p	F(1,32)	p_{adj}	Effect size η^2_p
Group	0.409	1.000	0.013	0.720	1.000	0.022
PhonProb	3.913	0.399	0.109	4.415	0.308	0.121
SylStr	2.032	0.820	0.06	3.902	0.308	0.109
Group x PhonProb	0.015	1.000	0.0005	0.013	1.000	0.0004
Group x SylStr	1.811	0.820	0.054	1.915	0.704	0.056
PhonProb x SylStr	0.116	1.000	0.004	0.225	1.000	0.007
Group x PhonProb x SylStr	3.464	0.432	0.098	4.347	0.308	0.120

Note: Data were transformed using the transformTukey() function from the rcompanion package, which determines a suitable transformation to maximize Shapiro-Wilk's W statistic. Here, data underwent a transformation of $x^{2.425}$.

Supplementary Table 11

Descriptive statistics and ANOVA assumptions for MMN mean amplitude of temporal deviants.

PhonProb	SylStr	Group	Descriptive Stats Mean (μV)	Std Dev	Shapiro-Wilk test W	p	Levene's test W	p	Box's M-test M	p
HPP	SylS1	DYS	-1.40	1.05	0.923	0.168	0.202	0.656	0.593	0.441
		CON	-1.67	1.05	0.919	0.145				
	SylS2	DYS	-1.75	1.19	0.994	1.00	1.63	0.211		
		CON	-0.932	0.894	0.947	0.409				
LPP	SylS1	DYS	-1.84	0.786	0.943	0.361	7.74	**0.009**		
		CON	-1.53	1.52	0.973	0.864				
	SylS2	DYS	-1.64	1.21	0.959	0.614	0.038	0.847		
		CON	-1.27	1.06	0.944	0.372				

Supplementary Table 12

2x2x2 split-plot ANOVA on MMN mean amplitude for temporal deviants (Bonferroni-Holm corrected)

Effect	$F_{(1,32)}$	p_{adj}	Effect size η^2_p
Group	1.750	1.000	0.052
PhonProb	1.087	1.000	0.033
SylStr	1.456	1.000	0.044
Group x PhonProb	0.073	1.000	0.02
Group x SylStr	2.581	0.826	0.075
PhonProb x SylStr	0.006	1.000	0.0002
Group x PhonProb x SylStr	1.506	1.000	0.045

Chapter 4

Phonological and temporal regularities lead to differential ERP effects in self- and externally triggered speech

Based on

ABSTRACT

Some theories of predictive processing propose reduced sensory and neural responses to anticipated events. Support comes from M/EEG studies, showing reduced auditory N1 and P2 responses to self- compared to externally generated events, or when stimulus properties are more predictable (e.g., prototypical). The current study examined the sensitivity of N1 and P2 responses to statistical regularities of speech. We employed a motor-to-auditory paradigm comparing ERP responses to externally and self-triggered pseudowords, varying in phonotactic probability and syllable stress. We expected to see N1 and P2 suppression for self-triggered compared to externally generated stimuli, with greater suppression effect for more predictable features such as high phonotactic probability and first syllable stress in pseudowords. We observed an interaction between phonotactic probability and condition on the N1 amplitude, with an enhanced effect of phonotactic probability in processing self-triggered stimuli. However, the directionality of this effect was reversed compared to what was expected, namely a larger N1 amplitude for high probability items, possibly indicating a perceptual bias toward the more predictable item. We further observed an effect of syllable stress on the P2 amplitude, with greater amplitudes in response to first syllable stress items, likely related to acoustic stimulus differences. The current results suggest that phonotactic probability may affect speech processing differently depending on whether it is self-triggered or externally presented. Further research is necessary to elucidate whether this can be related to feedforward models in speech production.

1   Introduction

The brain's capacity to formulate predictions of upcoming events in our environment is one of the most studied phenomena across sensory modalities (e.g., Baldeweg, 2006; Blakemore et al., 2000; Rao & Ballard, 1999). These predictions may relate to the timing ('when', temporal prediction) and content/quality ('what', formal prediction) of upcoming sensory events (Arnal & Giraud, 2012; Kotz & Schwartze, 2010), and are based on our acquired knowledge and experience of the world. A special form of prediction generated by the brain is related to the sensory consequences of our own actions. The underlying mechanism is described by the internal forward model of motor control (e.g., Wolpert & Miall, 1996). According to this model, when a motor plan is formulated, an internal copy of the command, termed "efference copy" is used to generate a prediction of the anticipated sensory feedback. This prediction, or "corollary discharge", is then compared to the actual sensory feedback (reafference signal), allowing the system to distinguish between self- and externally generated sensations, and to monitor and adapt our own motor output more readily. This model has also been applied to speech production, linking psycholinguistic models of feedback monitoring at the phoneme and syllable level to general motor control mechanisms (e.g., Hickok, 2012; Kotz & Schwartze, 2016)

As a consequence of this mechanism, the sensory response to internally generated stimulation is suppressed, leading to well-known phenomena such as the inability to tickle oneself (Blakemore et al., 2000). This perceived sensory suppression, termed motor-induced suppression (MIS), goes hand in hand with the suppression of sensory-related neural activity, shown across multiple sensory domains, including somatosensory (Blakemore et al., 2000) and auditory (e.g., Christoffels et al., 2011; Knolle et al., 2012; Niziolek et al., 2013). The degree of MIS reflects the accuracy of the generated prediction: the better the match between predicted feedback and actual sensory feedback, the greater the suppression. This pattern is supported by observations of an inverse relationship between noise level and suppression in fMRI: with higher noise levels, less suppression is observed (Christoffels et al., 2011). The magnitude of the neural activity in response to self-generated sensations is therefore thought to reflect the prediction error, or the mismatch between predicted and actual feedback: Noisy situations result in less clear and less predictable feedback, leading to less suppression (i.e., more neural activity).

MIS is modulated by stimulus properties, including the predictability of the frequency and timing of tones (Bäss et al., 2008; Knolle et al., 2013a) or manipulations of voice identity (Johnson et al., 2021), voice quality, and timing in speech (Aliu et al., 2009; Behroozmand & Larson, 2011). In a study comparing different utterances of the same vowel, self- or externally generated, Niziolek and colleagues (2013) observed greater suppression when the utterance showed formant ratios more prototypical for the individual speaker. Crucially, when the utterance deviated from the speaker's prototype, the auditory cortical response predicted the correction of the articulation. This observation confirms that this mechanism may be involved in monitoring and correcting behavior. MIS is further modulated by experience, with musicians showing different suppression patterns than non-musicians (Ott & Jäncke, 2013). In summary, these findings suggest that greater suppression is indicative of more predictable sensory events, and that this suppression may be modulated by experience.

These observations suggest that MIS may be a suitable measure to investigate the brain's sensitivity to regularities in the formal and temporal structure of speech during production. Within speech and language, regularities exist at multiple timescales, allowing the formulation of formal (e.g., phonotactic probability) and temporal (e.g., syllable stress) predictions across different processing levels. These predictions are established through exposure to regularities in speech throughout development, and evidence of sensitivity to these regularities is found already in infancy (Nazzi et al., 1998; Saffran et al., 1996). This sensitivity may provide an important foundation in the early stages of language acquisition, by allowing infants to segment the continuous speech signal into words (Jusczyk et al., 1999; Mattys & Jusczyk, 2001; Thiessen & Saffran, 2003), and continues to facilitate speech processing throughout the lifespan, as indicated by both behavioral and neural evidence.

Phonotactic probability modulates primarily sublexical language processes (i.e. independent of lexical/conceptual processing) in speech perception, such as nonword recognition (Luce & Large, 2001; Vitevitch & Luce, 1999) and recall (Thorn & Frankish, 2005). In contrast, lexical stress can guide the resolution of lexical conflict in spoken word recognition (e.g. <u>pre</u>sent = gift, pre<u>sent</u> = to give a presentation; Cutler, 2005). In production tasks, such as nonword repetition, both phonotactic probability (Edwards et al., 2004; Munson, Edwards, et al., 2005; Munson, Kurtz, et

al., 2005; Vitevitch & Luce, 1998, 2005) and syllable stress (Vitevitch et al., 1997) modulate performance, with high phonotactic probability items and more frequently occurring stress patterns (e.g. in Dutch, first syllable stress) being repeated more accurately. Furthermore, lexical stress can guide the learning of novel phonotactic constraints (Bian & Dell, 2020).

There is ample neural evidence supporting the aforementioned behavioral observations during speech perception, with variations in phonotactic probability and stress patterns modulations neural processing (Bonte et al., 2005; Di Liberto et al., 2019; Emmendorfer et al., 2020; Rothermich et al., 2012; Tremblay et al., 2016). However, data on the neural correlates of these features in speech production are sparse. fMRI investigations have shown sensitivity to distributional statistics such as phonotactic probability, syllable frequencies or mutual information in speech production tasks across the speech network, including auditory as well as motor regions, with reduced BOLD signal for items with higher frequency of occurrence within the language (Papoutsi et al., 2009; Tremblay et al., 2016). These findings are in line with psycholinguistic models proposing that motor plans of more frequently occurring structures are stored in a "mental syllabary", while less frequent articulatory representations need to be compiled from smaller units on the spot (Levelt, 1999; Levelt & Wheeldon, 1994; Schiller et al., 1996). Electrophysiological data on these features in speech production tasks is sparse. In a go/no-go task, where "go" decision was based on lexical stress position, N200 latency was earlier for words with first syllable stress (Schiller, 2006). However, this was proposed to be related to the incremental encoding (i.e. from word onset to end) of the meter during speech production, rather than a function of typical/atypical stress patterns, which is further supported by behavioral findings in trisyllabic stimuli (Schiller et al., 2006). Currently, we do not know of any studies investigating the effect of variations in phonotactic probabilities during speech production with electrophysiological methods.

The current experiment aimed to investigate how predictability of phonotactic probability and syllable stress contribute to speech production, extending our knowledge from previous studies investigating speech perception (e.g., Bonte et al., 2005; Emmendorfer et al., 2020) and production (e.g., Schiller, 2006; Tremblay et al., 2016). To approach this question, we focused our attention on motor-induced suppression, as this allows investigating how such (ir)regularities modulate the

accuracy of the prediction generated through the efference copy. While some studies have investigated this phenomenon in overt speech production (e.g., Aliu et al., 2009; Christoffels et al., 2011; Niziolek et al., 2013), this comes with challenges due to artifacts caused from engaging the facial muscles during articulation. Furthermore, overt production leads to variability in the pronunciation of the individual utterances, which can lead to changes in the degree of suppression (Niziolek et al., 2013). This is a particularly relevant constraint in the current design, as less familiar features may show more variability in articulation as well as more speech errors (Heisler & Goffman, 2016; Munson, 2001; Sasisekaran et al., 2010). To circumvent these challenges, we employed a button-press paradigm, or motor-to-auditory paradigm, where the participant elicits the presentation of speech stimuli via button-press (e.g., Knolle et al., 2013a, 2019; Ott & Jäncke, 2013; Pinheiro et al., 2018).

The classical design in these experiments employs three conditions: an auditory-only condition (AO), where participants are passively presented with auditory stimuli, a motor-auditory (MA) condition, where participants trigger the generation of self-produced pseudowords through a button-press, and finally a motor-only (MO) control condition used to correct for the motor component (MA – MO = MAC). This design has been applied to investigate MIS in response to a range of stimulus types, including tones (Knolle et al., 2013a), voices (Pinheiro et al., 2018), vowels (Knolle et al., 2019), and single syllables (Ott & Jäncke, 2013). These designs typically elicit modulations of the auditory N1 and P2 components. Observed reduction of N1 amplitude in response to self-triggered stimuli is thought to reflect an unconscious, automatic prediction resulting from the efference copy/corollary discharge, while P2 suppression reflects a more conscious differentiation between self- and externally generated events (e.g. Knolle et al., 2013a, 2019; Pinheiro et al., 2018). Here we investigated the effect of phonotactic and syllable stress regularities on MIS of the N1 and P2 components, using prerecorded utterances of bisyllabic Dutch pseudowords from each participant. Specifically, we aimed at testing the following hypotheses: (1) N1 and P2 amplitudes are reduced for self-triggered stimuli compared to externally generated stimuli (i.e. main effect of condition, MIS), (2) this reduction in amplitude is modulated by phonotactic probability and syllable stress (i.e. interactions between phonotactic probability and condition, and syllable stress and condition), with high phonotactic probability and first syllable stress items

leading to greater amplitude reduction due to greater predictability, and (3) phonotactic probability and syllable stress may interactively modulate motor-induced suppression (i.e. three-way interaction between phonotactic probability, syllable stress and condition), where we do not have precise predictions about the nature of this interaction.

2   Methods

2.1   Participants

34 right-handed native Dutch speakers participated in the study after giving their informed consent. The study was approved by the Ethical Committee of the Faculty of Psychology and Neuroscience at Maastricht University (ERCPN-OZL 205_17_03_2019) performed in accordance with the approved guidelines and the Declaration of Helsinki. Participants were invited to complete two sessions: one for recording the stimulus materials, followed by the EEG session. 5 participants completed the stimulus recording but did not complete the EEG session due to the COVID-19 pandemic. One participant was excluded from the EEG session due to failure to accurately reproduce stimuli. One participant was excluded due to excessive noise in the EEG signal (< 100 trials remaining per stimulus per condition). This led to a final sample of 27 participants (9 male, mean age: 21.9; standard deviation +/- 3.8), who completed both sessions of the experiment. The stimulus recording procedures and variations of the EEG paradigm were piloted in an additional 9 participants. The stimuli generated from these pilot participants were used to determine the criteria for stimulus selection as described in the following section.

Table 1

Stimuli

|	SylStr	
	SylS1	SylS2
PhonProb		
High (HPP)	**not**sal	not**sal**
Low (LPP)	**not**fal	not**fal**

Note. Bold font indicates stressed syllable (SylS1 = 1st syllable, SylS2 = 2nd syllable). PhonProb = phonotactic probability. SylStr = syllable stress. The phoneme combination '-ts-' constitutes the high phonotactic probability (HPP), and '-tf-' the low phonotactic probability (LPP).

2.2 Stimulus generation

The stimuli for the EEG experiment were prepared on an individual basis. Participants were invited for an initial stimulus recording session scheduled several days prior to the EEG session. The stimuli consisted of four pseudowords (Table 1), which differed from each other in phonotactic probability (notsal vs. notfal) and syllable stress (first vs. second syllable; adapted from Bonte et al., 2005; Emmendorfer et al., 2020). During the EEG experiment, each participant was presented with stimuli in their own voice. As second syllable stress is rare in Dutch, "natural" pronunciation of bisyllabic pseudowords with this stress pattern is challenging. To circumvent this issue, participants were presented with the target words, which were generated using a splicing procedure. The target words were spoken by a female Dutch speaker, who produced the syllables of interest by replacing them individually with syllables in existing bisyllabic Dutch words containing the same (spoken) consonant cluster and stress pattern as the target pseudowords (e.g. /**bad**zout/ → /**not**zout/ and /**bad**sal/ → **not**sal; /ont**slag**/ → /not**slag**/, and /ont**sal**/ → /not**sal**/; for more details see Emmendorfer et al., 2020). These spliced target words were presented to the participants of the current experiment. After ensuring the participant could hear and reproduce the differences between the pseudowords, each target was presented 15 times in random order, and the participants were asked to repeat them as accurately

as possible. Participants were not explicitly instructed to attend to the stress pattern as this could lead to exaggerated expression of syllable stress.

From the 15 repetitions of each pseudoword, one item was selected as the stimulus for the EEG experiment. To ensure comparability across participants, without having to manipulate the recording to deviate from the participants own naturally produced utterance, we selected items such that they were comparable in the timing of the perceptual centers (p-centers) of the syllables. P-centers are thought to represent the perceived "beat" of the speech stimulus. The timing of the p-centers was estimated with a beat detection algorithm (custom Matlab script adapted from Cummins & Port, 1998). Here, the beat, or p-center, is defined as the midpoint of each local rise in the amplitude envelope of the recorded signal, representing the vocalic nucleus of a syllable. The duration of the interval between the p-centers of each syllable in the bisyllabic pseudowords was calculated, and from 10 participants (9 pilot participants and 1 from the final sample), the average interval was calculated for each pseudoword. These values were used to select the best fitting stimulus for the participants who completed the subsequent EEG session. For each pseudoword, the item with the closest matching interval was selected. If this item contained acoustic artifacts or a mispronunciation, it was discarded, and the next best item was selected. This procedure allowed the selection of temporally comparable stimuli, while preserving each participant's own pronunciation without editing or manipulating the timing. A representation of the stimuli included in the experiment can be found in Figure 1. Stimuli were filtered with a Hann bandpass filter (80 – 10500 Hz), and intensity scaled to 60 dB. Mean stimulus duration was 0.640 s (standard deviation: 0.056 s), and the mean interval between p-centers of the stimuli was 0.319 s (standard deviation: 0.042 s).
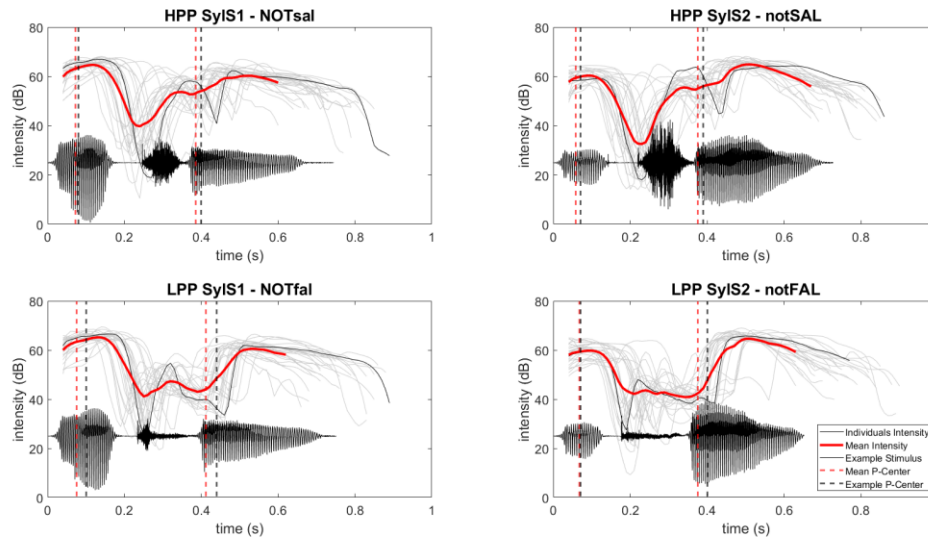
Figure 1: Stimuli. Stimuli selected for the EEG experiment. Individual intensity contours of the stimuli are represented in grey, mean intensity contours across participants in red. Stimuli from an exemplary participant are represented in black. Timing of the p-centers, representing the onset of the vocalic nucleus, are represented by dashed lines (red: averaged across participants, black: exemplary participant). Individual tokens to be used for each participant were selected based on the interval between the p-centers for first and second syllables. Stimulus onset t = 0 is equivalent to t = 0 in the subsequent ERP plots. HPP = high phonotactic probability, LPP = low phonotactic probability, SylS1 = first syllable stress, SylS2 = second syllable stress.

2.3   EEG paradigm

The paradigm (adapted from Johnson et al., 2021; Ott & Jäncke, 2013) consisted of three conditions (Figure 2A). In all three conditions, the trial began with the presentation of a fixation cross, followed by a cue (< left, > right) at 0.4 – 1.0 s after trial onset. In the motor-auditory condition (MA), participants pressed a button (left or right), which triggered the presentation of a stimulus. In the auditory-only condition (AO), participants were presented with the same cue, but the stimulus presentation occurred without button press, 0.5 s after cue onset. In the motor-only condition (MO), the participants pressed the cued button, but no stimulus was presented. This condition was included to correct for the motor component in the MA condition. This corrected motor-auditory condition (MAC) was calculated as MA – MO, thus allowing the comparison of neural activity in response to self-triggered (MAC) and externally generated auditory stimuli (AO). A reduction in N1 and P2 amplitudes for MAC relative to AO is then interpreted as motor-induced suppression.

The EEG recording occurred over the course of 6 experimental runs, each consisting of 18 blocks (8 MA, 8 AO, 2 MO) (Figure 2B). In each MA and AO block, one stimulus pair was presented. The stimuli within the pair differed from each other in either phonotactic probability or syllable stress (Figure 2C), and each cue/button

press corresponded to one stimulus. Each pair was presented twice per run and condition, with the cue/button assignment counterbalanced across blocks. Within each block, the first 4 trials (always including 2 left, 2 right) were excluded from analysis to allow the participant to form an association between cue and word. In four blocks per run (2 MA, 2 AO), four catch trials were included at the end of the block, where the cue-stimulus pairing was switched, i.e., the left cue was followed by the stimulus previously associated with the right cue. Participants were instructed to attend to the cue-stimulus pairing and were asked to report at the end of each block whether they noticed a switch. This task was included to ensure the participants were correctly associating the presented stimulus with the cue/button-press, and these trials were excluded from analysis. The total number of trials per block varied between 14 and 28 trials such that the participant could not anticipate when the catch trials would occur by counting. This resulted in 10 – 20 trials per block, and a total of 90 trials per condition/stimulus/cue assignment included in the analysis (Figure 2B).

2.4   EEG recording

EEG was recorded with BrainVision Recorder (Brain Products, Munich, Germany) using a 63-channel recording setup. Ag/AgCl sintered electrodes were mounted according to the 10% equidistant system, including 57 scalp electrodes, left and right mastoids for offline re-referencing, and four EOG electrodes to facilitate removal of artefacts caused by eye movements (2 placed on the outer canthi, 2 above and below the right eye). The scalp was cleaned at electrode sites and electrodes were filled with electrolyte gel to keep impedances below 10kΩ. Data was acquired with a sampling rate of 1000Hz, using Fpz as an online reference and AFz as ground. During recording, participants were seated on a comfortable chair in an acoustically and electrically shielded room.

2.5   EEG processing

EEG data was processed using the EEGLAB toolbox (Delorme & Makeig, 2004) and custom MATLAB scripts (MATLAB, 2018) The continuous EEG data were filtered using a bandpass filter of 1 – 30 Hz, and then downsampled to a sampling rate of 250 Hz. Noisy channels were identified, removed, and interpolated using the EEGLAB plugin clean_rawdata, and the data were re-referenced to the average signal of the

two mastoid electrodes. The data were then epoched 0 – 2.4 s relative to the onset of the trial to remove noisy break intervals, while still including the entire duration of the experimental blocks. The data were then decomposed using ICA. 2 – 4 independent components, reflecting blinks and horizontal eye-movements were removed for each participant.

Initial inspection of the reconstructed data, segmented to stimulus onset, revealed a positive deflection preceding the stimulus in all three conditions, which appears time-locked to the visual cue (Supplementary Figure S1). This deflection could not be removed through highpass filtering (Supplementary Figure S2) or ICA (Supplementary Figure S3). While the deflection was present in all three conditions, it is effectively removed during the MA – MO subtraction (Figure 3A). The remaining deflection in the AO condition likely reflects a combination of visual processing of the cue, as well as anticipatory processes and temporal orientation to the upcoming stimulus due to the fixed temporal interval between cue and stimulus in this condition (Figure 2A). This observation violates the assumption of baseline correction that there are no systematic differences between conditions in the selected window. Therefore, the data were instead baseline corrected to a window 0.2 s prior to the onset of the *cue,* and subsequently epoched -0.6 – 0.5 s relative to the onset of the *stimulus* or *button-press.*

Previous findings have shown differences between self- and externally triggered auditory stimuli already in the pre-stimulus window (Reznik et al., 2018), thus a pre-stimulus baseline window may not be appropriate for this type of paradigm even without the positivity we observe. The pre-cue baseline correction is used throughout the analysis steps. Before moving to the planned analysis of the N1 and P2 components, we first explored the pre-stimulus positivity through cluster-based permutation analysis.

Individual N1 and P2 peaks were manually determined from single subject average waveforms at electrode FCz for each stimulus and condition separately. While the auditory N1 and P2 are classically measured over the vertex electrode (Cz), we opted for a slightly more frontal site as visual inspection of the ERPs suggested the amplitudes to be less influenced by the pre-stimulus deflection at this channel (see Supplementary Figure S2). The N1 peak was determined as a negativity in the time window 100 – 300 ms following stimulus onset, P2 as a positivity following the N1

peak up until 400 ms. These time windows are later than the classically observed N1 and P2 windows, however, a relative delay is consistent with the nature of the stimuli due to their complexity (Conde et al., 2018) and slow onset rise time (Onishi & Davis, 1968). Furthermore, broad time windows were selected to determine the individual peak as we anticipated variability in their timing due to the variability of the individual stimuli (i.e., variations in rise time of first syllable between participants and between first and second syllable stress). If there was ambiguity in the selection of the peak within a waveform (e.g., 2 peaks within the given window), the peak with the more appropriate topography and timing relative to the participant's average as well as the grand average across participants was selected, to ensure the inclusion of comparable neural events across conditions and participants. The amplitude in a window +/-24 ms surrounding this latency was extracted for all scalp electrodes.
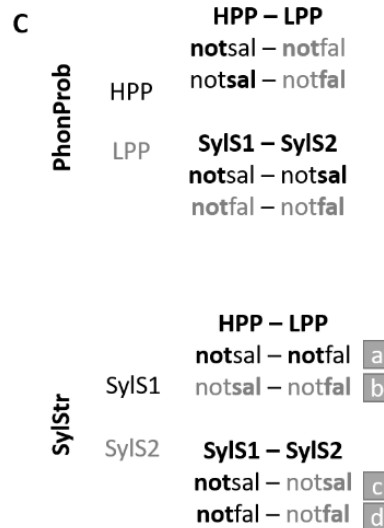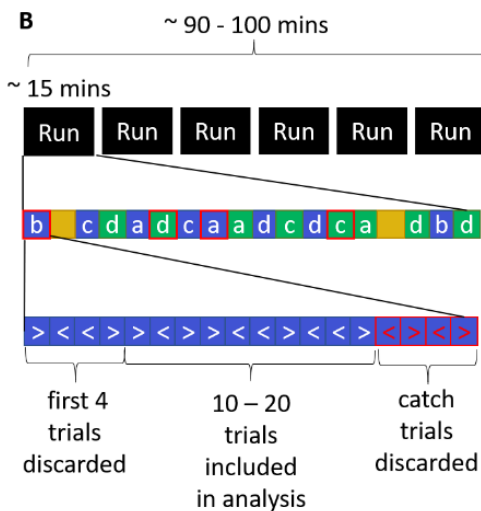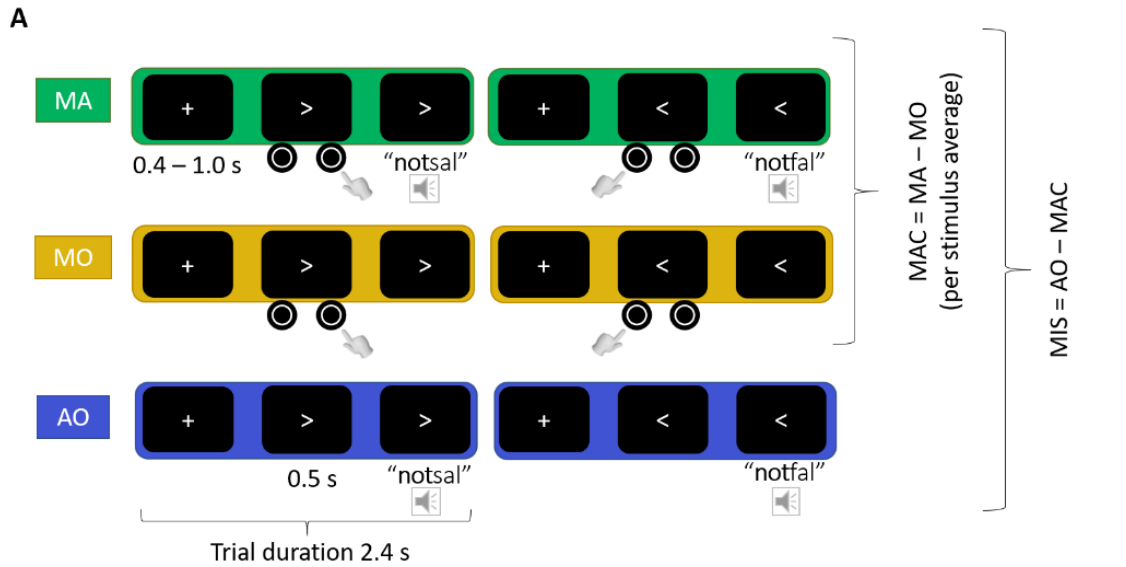
Figure 2: Experimental design. (A) Three experimental conditions. In all three conditions, the trial started with the presentation of a fixation cross for 0.4 – 1.0 s, followed by a visual cue < or >. In the MA (motor-auditory) condition, participants pressed the button corresponding to the cue (< left or > right), triggering the presentation of the pseudoword stimulus via speakers. In the MO (motor-only) condition, the participants pressed the button, but no stimulus was presented. In the AO (auditory-only) condition, the stimulus was presented 0.5 s after cue onset without button press. MA = motor-auditory, MO = motor only, AO = auditory only, MAC = motor-auditory corrected, MIS = motor-induced suppression (B) Overview of the EEG paradigm timeline. The total EEG measurement lasted 90 – 100 mins, consisting of 6 runs of approximately 15 mins each. Within each run, 18 mini-blocks were presented (8 AO, 8 MA, 2 MO). In each mini-block, one stimulus pair was presented (letters a-d correspond to the stimulus pair presented as denoted in panel C), where one stimulus was associated with the < (left) cue/button-press, one with the > (right) cue/button-press. Within the 8 mini-blocks of AO and MA, each pair was presented twice, with the cue/hand assignment counterbalanced across mini-blocks. AO and MA conditions consisted of 14 – 28 trials per mini-block. The first 4 trials were discarded from analysis. In four mini-blocks per run (2 AO, 2 MA), 4 catch trials where cue/hand assignment were switched were included at the end of a mini-block. These trials were also discarded from analysis, leading to a final 10 – 20 trials per mini-block. (C) Overview of stimuli and contrasted features: PhonProb = phonotactic probability, HPP = high phonotactic probability, LPP = low phonotactic probability, SylStr = syllable stress, SylS1 = first syllable stress, SylS2 = second syllable stress.

110
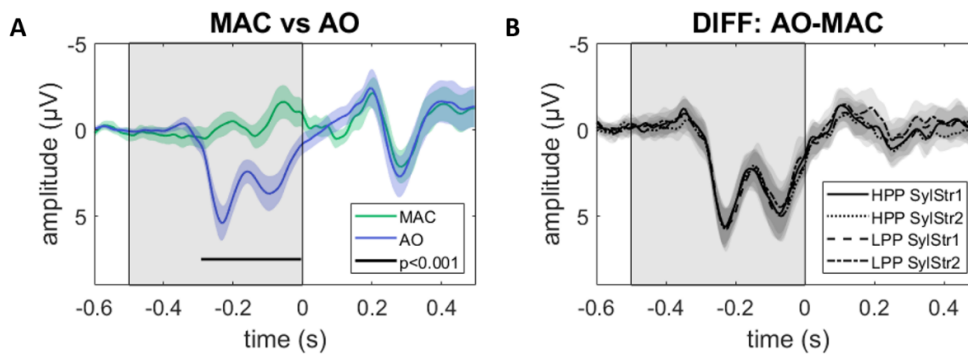
Figure 3. Overview of pre-stimulus deflection. (A) MAC (motor-auditory corrected, green) and AO (auditory only, blue) conditions, time-locked to stimulus onset and averaged across stimuli (+/-95% CI of the mean) in a frontocentral ROI (FCz, FC1, FC2, FC3, FC4). The shaded window indicates the window where cluster-based permutation tests were performed. The black line indicates the timing of the observed cluster with a significant difference between MAC and AO. (B) AO – MAC difference wave for individual stimuli at the same frontocentral ROI. The shaded area indicates the window where cluster-based permutation tests were performed to test for systematic effects of PhonProb, SylStr or an interaction thereof. No significant differences were found.

2.6   Statistical analyses

Due to the pre-stimulus deflection observed in the AO condition (Figure 3A), we first investigated whether this indeed reflected a systematic difference between AO and MAC. Such a systematic difference between conditions would render a direct comparison of the N1 and P2 amplitudes of these two conditions invalid, as we cannot exclude that any observed modulations of these components might be driven by this deflection rather than true motor-induced suppression as hypothesized. We tested this via a cluster-based permutation analysis (Maris & Oostenveld, 2007). One-sided paired-samples t-tests between AO and MAC were performed at each time-point in the time-window -0.5 - 0 s relative to stimulus onset for 1000 random partitions using the ft_timelockstatistics function of the Fieldtrip toolbox (Oostenveld et al., 2011). This analysis revealed a significant difference between AO and MAC. The observed cluster started at approximately 0.3 s prior to stimulus onset, with a broad topographic distribution. Based on this unexpected observation, a direct comparison of AO vs MAC at N1 and P2 components could not be interpreted as motor-induced suppression. Due to this finding, we were unable to test our specific hypotheses regarding the modulation of motor-induced suppression by phonotactic probability and syllable stress. However, as the broader aim of this research was to investigate the role of these regularities in speech production, we pursue analyses to answer the

question of how phonotactic probability and syllable stress modulate speech processing, and whether this differed across self- and externally generated speech.

Statistical analyses on N1 and P2 amplitudes were performed in R version 3.6.3 (R Core Team, 2013) using the rstatix package (Kassambara, 2019). Normal distribution of the N1 and P2 mean amplitude values was confirmed for all conditions via Shapiro-Wilk test (Supplementary Tables S1, S4), and outlier identification via boxplot methods did not reveal any extreme outliers (points beyond Q1 – 3*IQR, Q3 + 3*IQR). In a 2x2x2 (high vs. low PhonProb x first vs. second SylStr x AO vs. MAC Cond) repeated-measures ANOVA, we tested the following hypotheses for both N1 and P2 mean amplitudes averaged across electrodes in a frontocentral ROI (FCz, FC1, FC2, FC3, FC4): The N1/P2 amplitudes are modulated by the predictability of stimulus features (1) PhonProb and (2) SylStr, where more predictable utterances (i.e., high phonotactic probability and first syllable stress) lead to smaller amplitudes. Furthermore, (3) these features may interactively modulate N1/P2 amplitudes (PhonProb x SylStr interaction), and (4) may differ across conditions (PhonProb x Cond or SylStr x Cond interaction), where we would expect the MAC condition to show greater effects of these features due to error-monitoring. ANOVA results were corrected for multiple comparisons with Bonferroni-Holm correction using the adjust_pvalue() function (Cramer et al., 2016), and follow-up t-tests of simple effects were Bonferroni corrected.

3   Results

Visual inspection of the ERP grand averages (Figure 4A, 5A) reveals an N1/P2 morphology, with the N1 peaking around 200 ms and the P2 around 300 ms. When adjusted for the timing of the p-center of the first syllable of participants' pseudoword pronunciations, the N1 and P2 latencies are shorter, at approximately 125 and 212 ms, respectively. For our analyses we kept the time-locking to stimulus onset as it resulted in delayed but better aligned N1 and P2 responses across participants. In the following sections, we present the results of the statistical analyses. Here, we report only significant or otherwise noteworthy main effects and interactions, as well as post-hoc simple effects. The full results of the statistical analyses can be found in the supplementary materials (Supplementary Tables S1, S2 and S3 for N1, Tables S4 and S5 for P2 results).

3.1 N1

A 2x2x2 repeated measures ANOVA (high vs. low PhonProb x first vs. second SylStr x AO vs. MAC Cond) on N1 mean amplitudes (+/- 24 ms surrounding peak) averaged within the frontocentral ROI revealed a significant interaction between phonotactic probability and condition ($F(1,26) = 8.463$, p.adj = 0.049, $\eta^2_P = 0.246$). In the AO condition, LPP stimuli had a slightly larger N1 mean amplitude compared to HPP stimuli, while the reverse directionality was observed in the MAC condition (Figure 4A, B). This interaction was resolved by means of post-hoc paired samples t-tests testing the effect of phonotactic probability at each level of condition (Bonferroni corrected), which showed no significant effects for either AO ($t(26) = 0.840$, p.adj = 0.818, d = 0.162) or MAC ($t(26) = -2.04$, p.adj = 0.104, d = -0.392). Thus, the observed effect seems to reflect a crossover interaction, where the difference between HPP and LPP is different across conditions, but in neither AO or MAC do HPP and LPP differ from each other significantly. However, the effect size and mean amplitude difference is larger for MAC (d = -0.392, N1 mean amplitude HPP = -3.17 µV vs LPP = -2.67 µV) than AO (d = 0.162, N1 mean amplitude HPP = -2.87 µV vs LPP = -3.05 µV). No other main effects or interactions on N1 mean amplitude were significant (Figure 4B). We note that there was a small but significant difference in N1 latency across conditions ($t(26) = -2.43$, p = 0.022, d = -0.468), with the N1 peaking slightly earlier in the AO condition (mean latency = 190 ms) compared to MAC (mean latency = 197 ms).
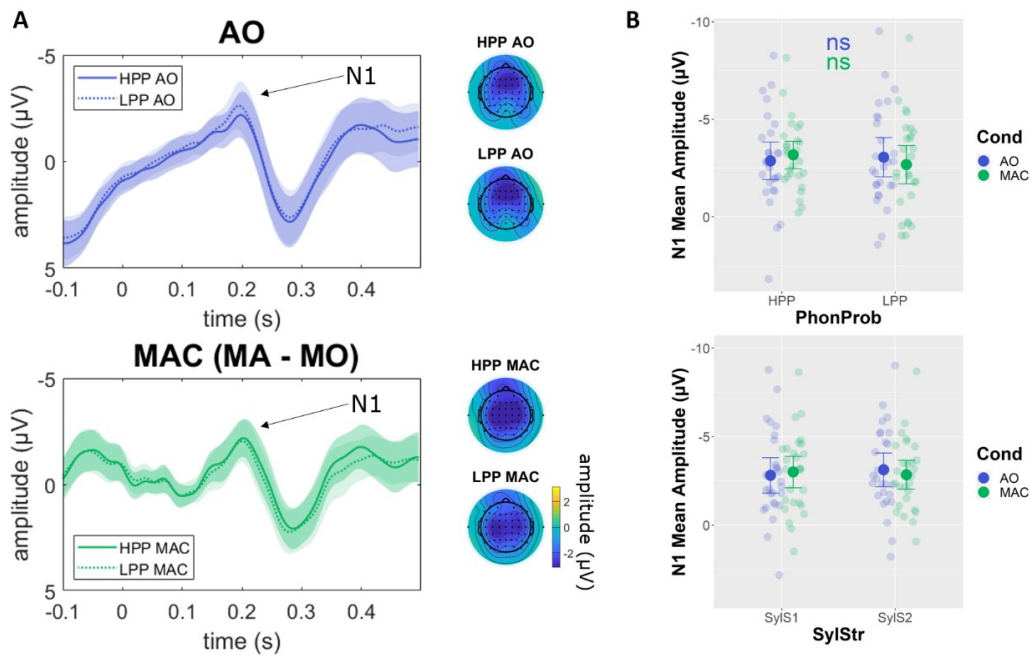
Figure 4: Interaction between phonotactic probability and condition on N1 amplitude. (A) Average ERP waveforms at frontocentral ROI (FCz, FC1, FC2, FC3, FC4) time-locked to stimulus onset, and corresponding topographies of N1 mean amplitudes. Note that baseline correction was performed in 200 ms window prior to *cue* onset (approximately -0.5 s). Due to the pre-stimulus positivity (see Figure 3), AO and MAC are different in the pre-stimulus window depicted here. (B) N1 mean amplitudes (+/- 24 ms surrounding individual peaks). HPP = high phonotactic probability, LPP = low phonotactic probability, AO = auditory only, MA = motor-auditory, MO = motor only, MAC = motor-auditory corrected, PhonProb = phonotactic probability, Cond = condition, ns = non-significant. Note: simple effects were not tested for SylStr x Cond as this interaction was not significant.
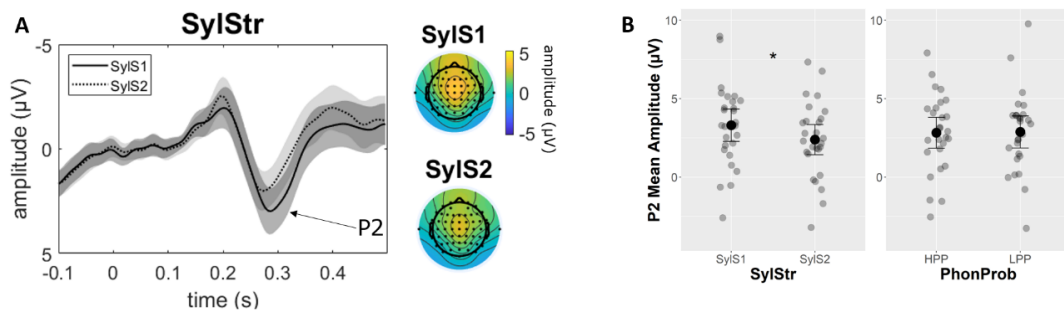
Figure 5: Main effect of syllable stress on P2 amplitude. (A) Average ERP waveforms at frontocentral ROI (FCz, FC1, FC2, FC3, FC4) time-locked to stimulus onset, averaged across phonotactic probability and condition, and corresponding topographies of individual P2 mean amplitudes (+/- 24 ms surrounding individual peaks). (B) P2 mean amplitudes (+/- 24 ms surrounding individual peaks). * p < 0.001. SylStr = syllable stress, SylS1 = first syllable stress, SylS2 = second syllable stress, PhonProb = phonotactic probability, HPP = high phonotactic probability, LPP = low phonotactic probability

3.2 P2

A 2x2x2 repeated measures ANOVA (high vs. low PhonProb x first vs. second SylStr x AO vs. MAC Cond) on P2 mean amplitude (+/-24 ms surrounding peak) averaged within the frontocentral ROI revealed a significant main effect of syllable stress $(F(1,26) = 22.993, p.adj < 0.001, \eta^2_P = 0.469)$. Stimuli with first syllable stress elicit a

114

larger P2 (mean amplitude = 3.30 μV) compared to those with second syllable stress (mean amplitude = 2.39 μV; Figure 5). No other main effects or interactions were significant.

4   Discussion

The current study aimed to investigate whether motor-induced suppression of the N1 and P2 amplitudes is modulated by formal (phonotactic probability) and temporal (syllable stress) predictability in the speech signal. We used a motor-to-auditory paradigm, where participants triggered the generation of self-produced pseudowords through a button-press. This approach was intended as a step towards investigating speech production, while limiting the interference of motor artifacts and speech errors present during overt production of pseudowords. We expected to observe a motor-induced suppression effect, with larger N1 and P2 amplitudes in the auditory-only condition, compared to the motor-auditory condition. Furthermore, we expected this suppression effect to be modulated by phonotactic probability and/or syllable stress, where high probability items (high phonotactic probability and first syllable stress) would elicit greater suppression, as they might be more "prototypical" items in the language (Niziolek et al., 2013). Due to an observed pre-stimulus deflection in the auditory-only condition, not present in the motor-auditory condition after correcting for motor output (Figure 3), we were not able to test our specific a priori hypotheses regarding motor-induced suppression in the current data. However, the design still allowed investigating the broader question as to how phonotactic probability and syllable stress contribute to the processing of self- and externally generated speech. We observed a modulation of N1 amplitude by phonotactic probability, which was enhanced in response to self-triggered stimuli, and a modulation of P2 amplitude by syllable stress, with first syllable stress eliciting a larger P2 compared to second syllable stress. While these analyses are post-hoc in nature, they provide insights into the differences in processing phonotactic and temporal regularities in self- and externally produced speech that can be followed up upon in future experimental designs.

We investigated the effect of variations of phonotactic probability and syllable stress across self- and externally generated conditions on N1 mean amplitude. Here, we observed a significant interaction between phonotactic probability and condition,

but no significant main effects or simple effects. Inspecting the data revealed that this interaction was a crossover effect, indicating that the direction of the effect of phonotactic probability differs across conditions (i.e., HPP > LPP in MAC, LPP > HPP in AO). However, the simple effects did not reach significance in either the auditory-only or motor-auditory condition. It is noteworthy though that visual inspection of amplitudes as well as comparison of effect sizes revealed the difference between high and low phonotactic probability to be larger for self-triggered words compared to externally generated ones. Thus, the pattern we observed in this interaction is in line with the notion that such regularities have greater weight in speech production due to feedforward processes. Variations in phonotactic probability of planned utterances may require different degrees of monitoring, as they may be more or less likely to result in mispronunciation. Indeed, it has been shown that phonotactic probability modulates accuracy and speed in speech production (Edwards et al., 2004; Munson, Edwards, et al., 2005; Munson, Kurtz, et al., 2005; Vitevitch & Luce, 2005). However, we hypothesized a larger amplitude for the less probable item as this would generate greater surprise, while our data suggest the opposite pattern. One could assume that these effects might be driven by differences in the motor-only (MO) condition, however we can exclude this possibility, as this condition does not include variations in stimulus type (i.e., same trials of MO are subtracted from all MA averages).

A closer look at theories of predictive processing may explain this discrepancy. While cancellation theories predict suppression of predicted sensory events to highlight novel or unexpected events (e.g., Blakemore et al., 2000), Bayesian theories suggest a perceptual bias or gain for predictable events (e.g., de Lange et al., 2018). Recent developments propose a two-process model to resolve the conflict between these contradictory theories (Press et al., 2020). Here it is proposed that the perceptual system is tuned toward expected events when there is a large overlap between prior and posterior probabilities (i.e., low surprise), resulting in perceptual gain for expected events, as suggested by Bayesian accounts (e.g., Thomas et al., 2020). When there is little overlap between these prior and posterior distributions (i.e., high surprise), this suggests that the model of the environment must be updated, resulting in higher activation to signal the unexpected event, in line with cancellation theories.

The pseudowords used in our design all consist of legal phonotactic structures and lexical stress patterns. Furthermore, the trials included in the analysis did not include any violations of predicted stimuli. Thus, the surprise generated by any given stimulus was low, and would not require the system to update their model of the world. Instead, given the acoustic similarity of high and low phonotactic stimuli, it is more likely that perception was biased toward the high probability item *notsal*. The trend toward a larger N1 amplitude for high probability items in the motor-auditory condition supports this notion. Interestingly, the timing of this modulation around 200 ms suggests that it may occur prior to, or concurrently with, the actual manipulation of phonotactic probability, which occurs at the syllable boundary (occurring around 200 – 250 ms, see Figure 1). This is in line with the opposing process theory proposed by Press and colleagues, as effects of perceptual sharpening are often observed prior to or within 50 ms of the expected stimulus, preceding cancellation effects (e.g., Press & Yon, 2019; Yon & Press, 2017). The perceptual sharpening may also render the system more sensitive to coarticulatory cues already present within the first syllable.

We did not observe a comparable modulation of the N1 by syllable stress, despite this feature also varying in probability in the Dutch language, with first syllable stress being the more common pattern in bisyllabic words. This observation is in line with a previous study in speech perception (Emmendorfer et al., 2020), where variations in syllable stress did not modulate MMN amplitudes. A range of other studies conducted in languages with a fixed stress pattern, such as Hungarian (Honbolygó & Csépe, 2013) or Finnish (Ylinen et al., 2009) does however show a modulation that is in line with a violation response (i.e. larger MMN to the illegal stress pattern). The divergent results here indicate that Dutch speakers process variations in lexical stress patterns differently than their Hungarian or Finnish speaking counterparts. While syllable stress may be exploited during development (Weber et al., 2004), and continues to play a role in resolving lexical conflict in spoken word recognition (Cutler & Van Donselaar, 2001), predictions generated relating to stress patterns may be weaker compared to those relating to phonotactic probability, particularly in the case of pseudoword stimuli in languages without a fixed lexical stress pattern.

We further investigated the effect of variations of phonotactic probability and syllable stress across self- and externally generated conditions on P2 mean amplitude. Here, we observe a main effect of syllable stress, where first syllable stress stimuli elicit a larger P2 amplitude compared to second syllable stress stimuli. This observation may be explained by acoustic differences in first and second syllable stress items (see Figure 1). The main perceptual markers of lexical stress are intensity, pitch, and duration of the syllable. Thus, while the stimuli were equalized in intensity across the whole word, they differed in the first syllable, with first syllable stress items having a greater intensity compared to second syllable stress items. The timing of our P2 at around 280 ms suggests that this component reflects information from the first syllable. We do not observe an interaction with condition, which would suggest a conscious differentiation between self- and externally generated events (Knolle et al., 2013a, 2019; Pinheiro et al., 2018). Therefore, this pattern is more in line with previous observations of P2 amplitude being modulated by stimulus intensity (for review, see Crowley & Colrain, 2004). However, if this is indeed a purely acoustic effect, one would expect to find a similar modulation in the N1 component, which we do not observe. It is possible that the amplitude modulation from the syllable stress pattern is masked due to distortion of the overall N1 amplitude from overlap of the pre-stimulus deflection. However, assuming that this distortion is equal across stimuli we would still expect to observe an effect in the N1 amplitude across first and second syllable stress stimuli.

An alternative explanation for the P2 modulation may lie in categorical perception of speech. The neural correlates of categorical perception around the typical P2 time-window, indicated by investigations of phoneme processing comparing tokens that vary along a continuum across phoneme boundaries (e.g., Bidelman et al., 2013, 2020). Here, ambiguous speech sounds show a smaller P2 amplitude compared to speech sounds that clearly fall within a phoneme category. As second syllable stress is atypical for Dutch bisyllabic words, this may come with variability in the pronunciation, including variability in vowel quality. However, as the current experiment did not specifically modulate categorization, and we also do not have data on the perceptual categorization of the vowels in the current design, the current data cannot address this question completely, thus this interpretation remains speculative.

The temporal dissociation of the observed effects, with phonotactic probability modulating N1 and syllable stress P2 amplitude, may suggest differences in the time course of processing of these features. In the current design, the features phonotactic probability and syllable stress are manipulated at different points in time of the stimuli: first and second syllable stress items differ from each other in principle from stimulus onset on, while phonotactic probability is varied at the syllable boundary. Furthermore, as the stimuli are naturally produced, we have little control over the precise timing of acoustic markers of phonotactic probability and syllable stress relative to the timing of the ERP components of interest. This makes it difficult to disentangle differences in the time course of the neural processing of these linguistic features from differences in when the information relevant to these features becomes available in the specific pseudoword stimuli. Thus, it is not possible to draw general conclusions about the relative time course of processing of phonotactic probability and syllable stress beyond the current design.

Although we originally set out to test hypotheses relating to motor-induced suppression, limitations to the current design hinder us from following the original analysis plan. The pre-stimulus deflection observed in the auditory-only condition (Figure 3) draws attention to the cue as a confound. This deflection appears to be time-locked to the cue onset, covers a broad time window and is present across the scalp, though larger in amplitude at more parietal regions. While the cue is identical in all conditions, it is effectively subtracted out from the motor-auditory condition along with the motor component (MAC = MA - MO) but remains present in the auditory-only condition. Including a visual control to subtract from the auditory-only condition may ameliorate this issue, however this would only account for purely visual processes. The deflection likely also represents attentional and anticipatory processes, as the participant was instructed to explicitly attend to the stimulus and could anticipate not only *which item* would be presented, but also *when* it would be presented, due to the constant timing between cue and stimulus. Thus, an additional adjustment to the current paradigm could include jittering the timing of these events to dissociate the processes associated with the cue and the stimulus. Varying the time between cue and stimulus could also address the question of whether the suppression effect is driven by the temporal predictability of the stimulus (Hughes et al., 2013; Sowman et al., 2012).

We do not observe any suppression in either the N1 or P2 components for the motor-auditory vs auditory-only condition. This observation differs from the bulk of similar studies applying this type of paradigm comparing the processing of self- and externally-generated stimuli (e.g., Bäss et al., 2008; Knolle et al., 2012, 2013a; Pinheiro et al., 2018). The lack of suppression may be explained by variations in the design between previous studies and the current experiment. The typical approach in this paradigm does not include a cue. Instead, the paradigm is typically applied as a blocked design (but see Knolle et al., 2013b for an event-related variation), where the button-presses generating the stimulus presentation are self-initiated in the motor-auditory condition. The auditory stimuli are then presented at the same temporal intervals in the auditory-only condition, again without a cue. Thus, a crucial difference between the auditory-only and motor-auditory conditions in these approaches, in addition to whether the stimulus is self-triggered or externally presented, is the predictability of the stimulus timing: in the motor-auditory condition, the participant can accurately predict the timing, while some temporal uncertainty remains in the auditory-only condition. A considerable portion of the suppression effect observed in previous research may therefore be driven by the temporal predictability of the events (Hughes et al., 2013; Sowman et al., 2012). In the current study, the timing of the stimulus in the auditory-only condition is predictable due to the cue, thus this difference between the auditory-only and motor-auditory conditions does not exist. If temporal predictability indeed drives the suppression effect, it is therefore not surprising that we do not observe this effect in the current paradigm. Future studies investigating the suppression effect should therefore consider not only varying the formal predictability of the stimulus, but also its temporal predictability.

In conclusion, the present experiment provides preliminary insights into differences in processing phonotactic and temporal regularities in speech perception and production, by comparing self-triggered (via button press) to externally generated (own) speech. Our findings suggest that phonotactic regularities play an important role in processing self-triggered speech, with a perceptual bias toward more probable phonotactic structures, in line with Bayesian accounts of predictive processing, or a combined model incorporating both Bayesian and cancellation theories. We further observe an effect of syllable stress, which is likely explained by the acoustic differences between the first syllable in pseudowords with first and

second syllable stress. To summarize, the current research suggests that a sensitivity to regularities in phonotactic and temporal structure of speech may be differently exploited in speech perception and production processes. Further investigations controlling for some of the limitations observed in the current paradigm are needed to confirm the results of the current post-hoc analyses.

Author Contributions: AKE, MB, BMJ and SAK designed the experiment, AKE prepared materials, collected and analyzed the data, and wrote the first draft of the manuscript, AKE, MB, BMJ and SAK refined the manuscript.

Supplementary Materials
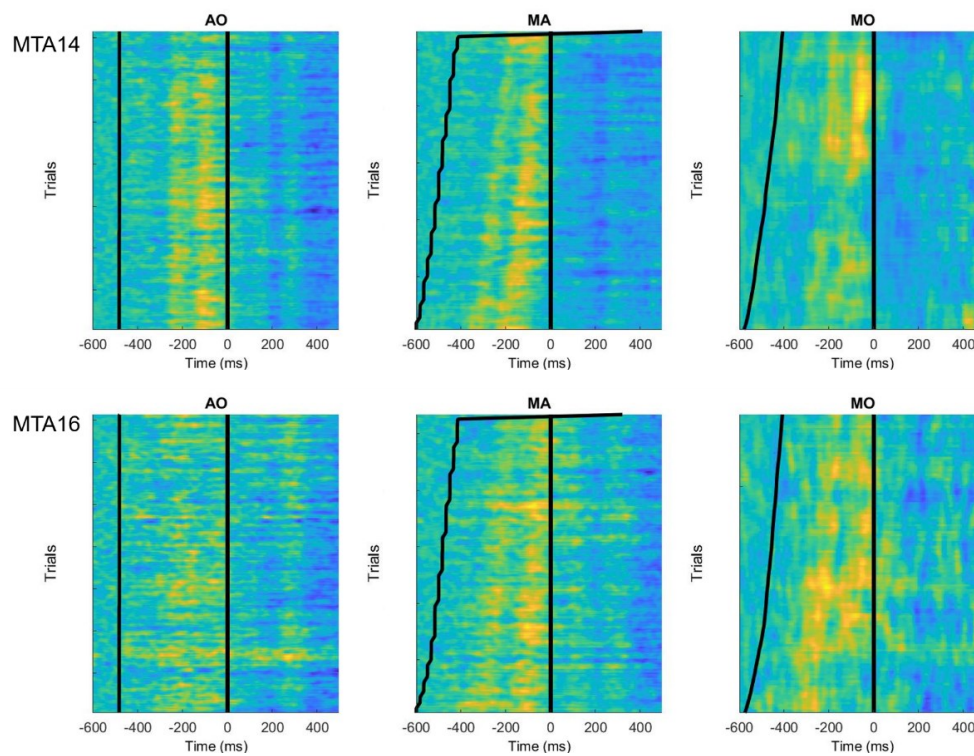
Figure S1: Exemplary participants MTA14 and MTA16 illustrating pre-stimulus deflection is related to visual cue. Single trial data at channel FCz represented on the y-axis, color representing amplitude in arbitrary scale (yellow = positive, blue = negative). In MA and MO conditions, trials are sorted based on interval between cue (black line at around t = -500 ms) and stimulus (black line at t = 0). Pre-stimulus positivity is aligned with timing of cue in both MA and MO condition.
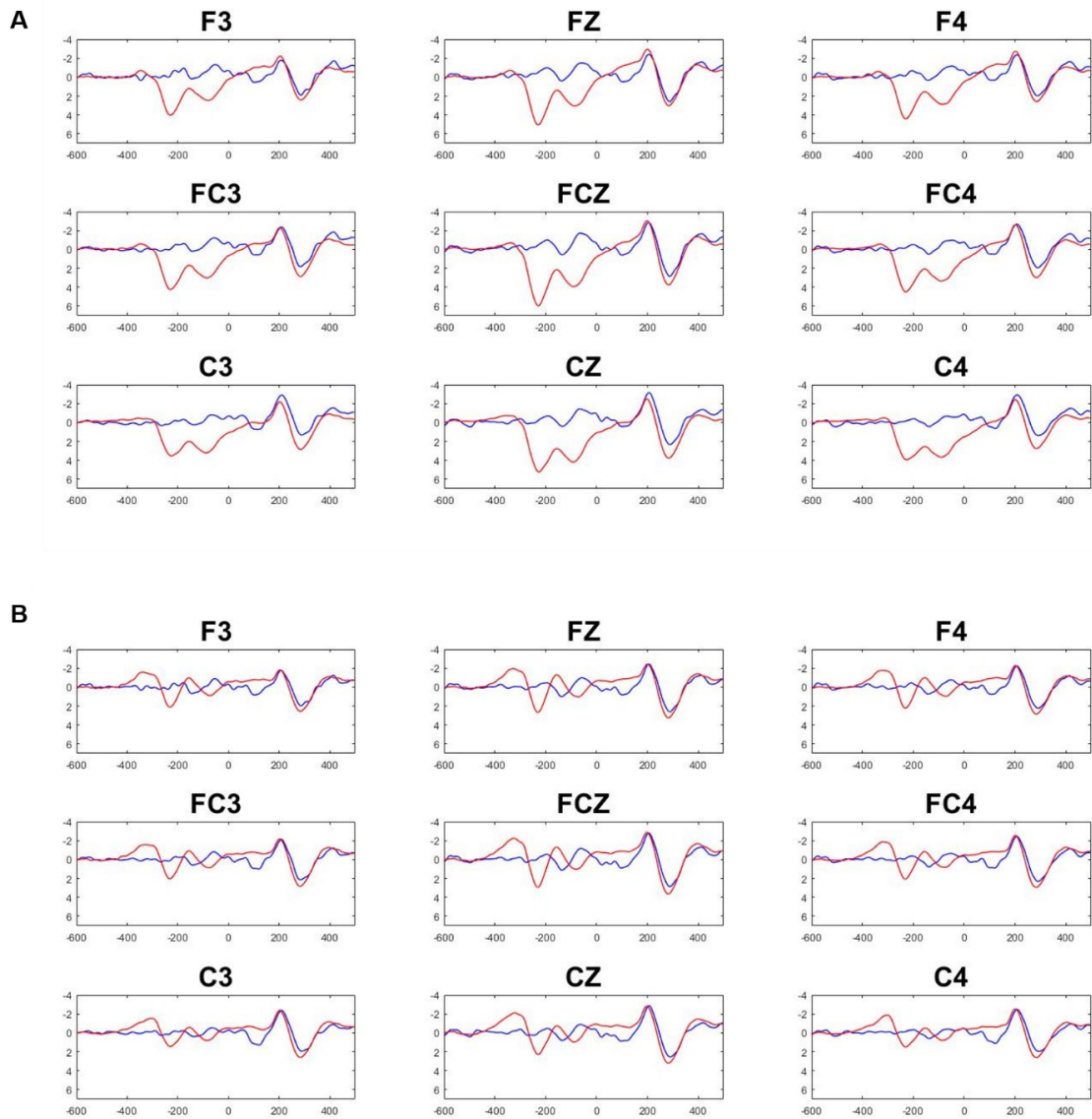
Figure S2: AO vs MAC -600 to 500 ms relative to stimulus onset in sample of n = 22. (A) 1-30Hz filter used in final analyses (B) 3-30Hz filter. Higher high-pass cutoff reduces overall amplitude of pre-stimulus deflection but does not remove it.
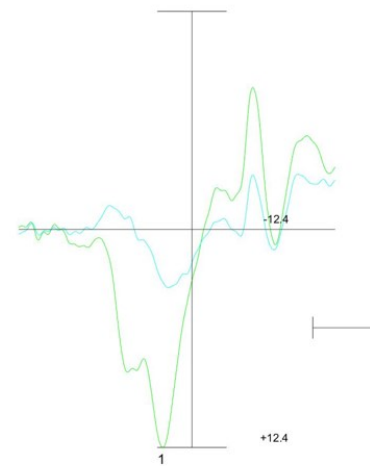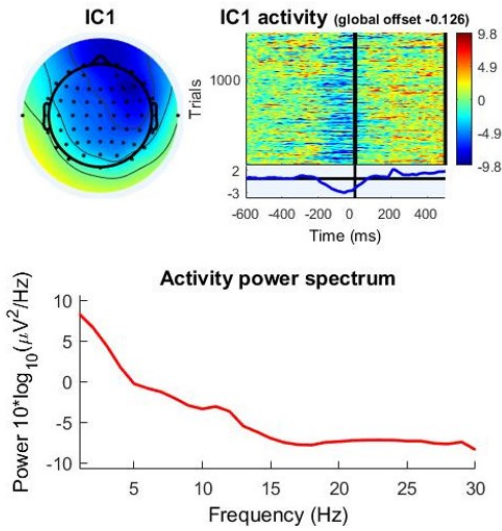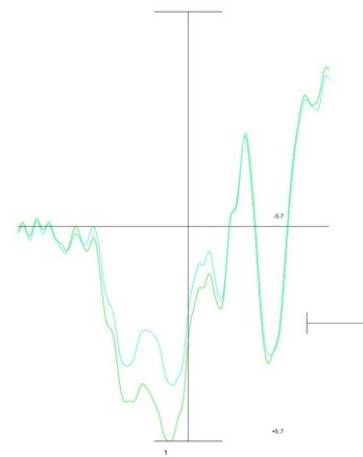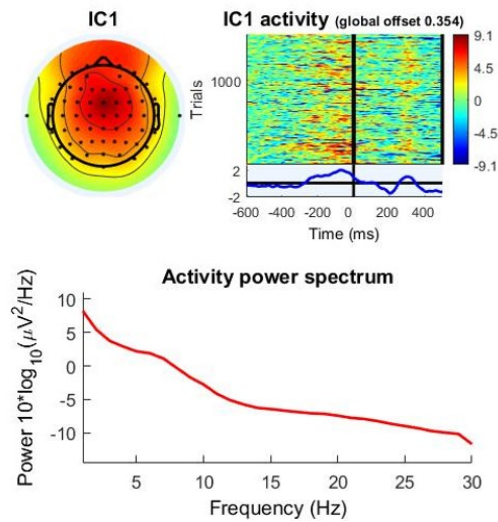
Figure S3: Exemplary participants MTA14 and MTA16: ICA fails to separate pre-stimulus deflection from components of interest. Left: topography, time course and power spectrum of independent component (IC) containing pre-stimulus deflection. The same component also includes deflection in the time window of N1 and P2 effects. Right: grand average ERP at channel FCz across all conditions. Green includes all ICs, blue with the selected IC removed. Removing this component often substantially reduces N100 amplitude (MTA14) or does not affect the positivity as desired (MTA16).

Supplementary Table S1
Descriptive statistics N1 mean amplitudes

Within-subjects variables			N1 mean amplitude		Shapiro-Wilk test	
Cond	**PhonProb**	**SylStr**	**Mean**	**Std Dev**	**W**	**p**
AO	**HPP**	**SylS1**	-2.94	2.48	0.967	0.536
		SylS2	-2.80	2.58	0.983	0.919
	LPP	**SylS1**	-2.67	2.75	0.934	0.087
		SylS2	-3.44	2.52	0.957	0.319
MAC	**HPP**	**SylS1**	-3.32	2.01	0.980	0.853
		SylS2	-3.02	1.82	0.957	0.314
	LPP	**SylS1**	-2.68	2.66	0.981	0.891
		SylS2	-2.67	2.51	0.976	0.763

Supplementary Table S2
2x2x2 Repeated measures ANOVA on N1 mean amplitude (Bonferroni-Holm corrected)

Effect	F(1,26)	p_{adj}	Effect size η^2_p
PhonProb	0.605	1.000	0.023
SylStr	0.238	1.000	0.009
Cond	0.010	1.000	<0.001
PhonProb x SylStr	4.222	0.300	0.140
PhonProb x Cond	8.463	0.049 *	0.246
SylStr x Cond	3.541	0.355	0.120
PhonProb x SylStr x Cond	2.624	0.468	0.092

Supplementary Table S3
Post-hoc paired samples t-tests of effect of PhonProb at individual levels of Condition. (HPP vs LPP)

Condition	t(26)	p_{adj}	Effect size (d)
AO	0.840	0.818	0.162
MAC	-2.04	0.104	-0.392

Supplementary Table S4
Descriptive statistics P2 mean amplitudes

Within-subjects variables			N1 mean amplitude		Shapiro-Wilk test	
Cond	PhonProb	SylStr	Mean	Std Dev	W	p
AO	HPP	SylS1	3.56	3.42	0.988	0.984
		SylS2	2.70	3.00	0.973	0.688
	LPP	SylS1	3.46	3.35	0.961	0.387
		SylS2	2.37	3.29	0.979	0.834
MAC	HPP	SylS1	3.05	2.60	0.966	0.493
		SylS2	1.96	2.18	0.979	0.828
	LPP	SylS1	3.14	2.69	0.975	0.741
		SylS2	2.53	2.43	0.970	0.596

Supplementary Table S5
2x2x2 Repeated measures ANOVA on P2 mean amplitude (Bonferroni-Holm corrected)

Effect	$F_{(1,26)}$	p_{adj}	Effect size η^2_p
PhonProb	0.065	1.000	0.003
SylStr	22.993	0.0004***	0.469
Cond	0.760	1.000	0.028
PhonProb x SylStr	0.092	1.000	0.004
PhonProb x Cond	4.292	0.288	0.142
SylStr x Cond	0.232	1.000	0.009
PhonProb x SylStr x Cond	1.508	1.000	0.055

Chapter 5

Summary and General Discussion

The brain makes use of regularities in the sensory environment to formulate predictions that facilitate perception. Such predictive capacities have been proposed to influence neural processing across multiple domains, including speech and language processing. This dissertation aimed to investigate the electrophysiological correlates of phonological and temporal regularities in speech processing. In three empirical studies, we investigated the neural sensitivity to variations in phonotactic probability (phonological regularity) and syllable stress patterns (temporal regularity) in native Dutch speakers. We first developed a passive oddball paradigm where we manipulated these features simultaneously in typically reading adults (Chapter 2). This paradigm was then applied to dyslexic readers, to investigate how they differ in their sensitivity to these features (Chapter 3). Here, we also investigated how the time-frequency correlates of the mismatch response differ between variations in phonotactic probability and syllable stress. Finally, we made a step towards speech production, where we tested the sensitivity of the N1 and P2 ERP components to phonological and temporal regularities in self- and externally triggered speech (Chapter 4). The following section contains a summary and discussion of the findings presented in this dissertation.

1   Summary

Current theories of neural processing propose that the brain formulates predictions of upcoming sensory events to facilitate perception (Engel et al., 2001; Raichle, 2010; Rao & Ballard, 1999). Such predictions are based on learned regularities in the environment, and can be formed about both the content and the timing of external sensory input (Arnal & Giraud, 2012), as well as the sensory consequences of our own actions (Wolpert & Miall, 1996). These principles also apply to speech processing. The speech signal is highly complex and variable, and yet listeners are able to understand speech even when sensory input is suboptimal (e.g., Hannemann et al., 2007). Predictive processing has been proposed as a mechanism to facilitate this seemingly effortless processing at multiple levels of linguistic analysis (e.g., DeLong et al., 2005; Freunberger & Roehm, 2016). Even in the absence of semantic or lexical content, our brain is tuned to regularities in the speech signal (Bonte et al., 2005; Emmendorfer et al., 2020; Vidal et al., 2019). In any given language, certain combinations of speech sounds or stress patterns are more likely to occur than others, and sensitivity to such

sublexical regularities can guide perception and facilitate production (Edwards et al., 2004; Luce & Large, 2001; Munson, 2001; Vitevitch et al., 1997; Vitevitch & Luce, 2005). In three empirical chapters, this dissertation investigated predictive processing of two of these types of regularities, phonotactic probability and syllable stress, using EEG.

In Chapter 2, we set out to investigate the neural sensitivity to variations in phonotactic probability and syllable stress using the identity mismatch negativity (Eulitz & Lahiri, 2004) in a passive oddball paradigm. Prior research in native Dutch speakers has shown that the MMN is modulated by variations in phonotactic probability in both children (Bonte et al., 2007) and adults (Bonte et al., 2005; Noordenbos et al., 2013). In these studies, deviants with high phonotactic probability elicited a larger MMN compared to deviants with low phonotactic probability, indicative of facilitated perceptual change detection for high probability items. MMN sensitivity to syllable stress variations has been demonstrated in fixed-stress languages such as Finnish (Ylinen et al., 2009) or Hungarian (Honbolygó et al., 2004; Honbolygó & Csépe, 2013; Ragó et al., 2014), however studies in languages with variable stress patterns such as Dutch are sparse and often focus on the MMN sensitivity to acoustic markers of stress rather than more abstract sensitivity to variations in stress patterns (e.g., Tong et al., 2014; Zora et al., 2015). The current approach aimed to, for the first time, investigate MMN sensitivity to phonotactic probability and syllable stress patterns when they are manipulated simultaneously, as these features also vary simultaneously in natural speech. Our findings support previous observations of MMN sensitivity to variations in phonotactic probability (Bonte et al., 2005; Noordenbos et al., 2013), with high phonotactic probability deviants eliciting an earlier MMN compared to low phonotactic probability deviants. No modulation of the MMN by variations in syllable stress was found, suggesting that Dutch speakers may not be sensitive to this feature when it is manipulated in pseudowords. While we did not observe any direct interactions between phonotactic probability and syllable stress, the current modulation pattern (latency effect) differed from prior observations with similar stimuli (amplitude effect; Bonte et al., 2005). A possible explanation for this discrepancy might be that the simultaneous manipulation of syllable stress changes the processes underlying the amplitude effect observed in previous studies.

132

In Chapter 3, the oddball paradigm first implemented in Chapter 2 was applied to dyslexic readers. Here, we aimed to investigate whether individuals with dyslexia show altered sensitivity to phonotactic probability and syllable stress. Reduced sensitivity to variations in phonotactic probability has been observed in both children (Bonte et al., 2007) and adults (Noordenbos et al., 2013) with dyslexia. Further evidence suggests that children who later develop reading impairments show atypical MMN responses to variations in syllable stress (Schaadt & Männel, 2019), in line with frameworks suggesting that impaired encoding of speech rhythm may be an underlying factor in developmental dyslexia (Goswami, 2011; Ladányi, Persici et al., 2020; Lallier et al., 2017). Here we aimed to replicate previous findings on phonotactic probability in an adapted paradigm where it was manipulated simultaneously in syllable stress, and for the first time investigated the sensitivity of Dutch adults with dyslexia to variations in syllable stress. Additionally, using time-frequency analysis, we investigate the oscillatory correlates of phonological and temporal change detection. Both formal and temporal deviants elicit a significant MMN, indicating auditory change detection. Our MMN findings support previous observations of reduced sensitivity to phonotactic probability in dyslexic adults (Noordenbos et al., 2013). We did not observe any MMN modulation of syllable stress in dyslexics or controls. The distinct oscillatory correlates of the mismatch negativity for processing formal and temporal deviants are a particularly interesting finding in this chapter. Deviants differing from the standard in phonotactic probability (formal deviants) exhibit the typical increased theta ITC associated with sensory memory processes underlying deviancy detection (Fuentemilla et al., 2008; Hsiao et al., 2009), which was enhanced in dyslexic readers, in line with previous observations (Halliday et al., 2014). This enhanced phase-locking to deviant stimuli in dyslexic readers may reflect an involuntary attention switch to evaluate the deviant stimulus (Fitzgerald & Todd, 2020; Tamura et al., 2015), suggesting atypical auditory interference control (Gabay et al., 2020). Deviants differing in syllable stress (temporal deviants) on the other hand showed a decrease in delta/theta ITC and did not differ between dyslexic and typical readers. This response pattern likely suggests a disruption in phase-locking to the temporal deviant relative to the standard, due to the change in the temporal structure of the stimulus when the syllable stress pattern is altered. Taken together, these time-frequency results suggest dissociable oscillatory mechanisms

underlying the mismatch response for formal and temporal deviants. We observed atypical processing of formal deviants in dyslexics, in line with prior reports (Bonte et al., 2007; Halliday et al., 2014; Noordenbos et al., 2013), however, no differences between dyslexic and typical readers in processing temporal deviants. Some reports have highlighted that the difficulties associated with stress perception may rather be related to explicit stress awareness, or task demands when stress perception is tested with an explicit task (Anastasiou & Protopapas, 2015; Barry et al., 2012; Mundy & Carroll, 2012, 2013). Thus, dyslexic readers may not differ from controls when this sensitivity is tested implicitly as in the passive oddball paradigm.

In Chapter 4, we aimed to test whether predictability of phonotactic and stress regularities are exploited differently in speech perception and production. For this, we applied a motor-to-auditory paradigm, where participants listen to stimuli that were either self-triggered via button press, or externally generated. In such a paradigm, self-triggered stimuli lead to a suppressed neural response compared to identical stimuli that are externally presented, where the magnitude of this suppression, termed motor-induced suppression (MIS), depends on the predictability of the stimuli (e.g., Bäss et al., 2008; Knolle et al., 2013a). While such paradigms have been applied using tones (Knolle et al., 2013a), vowels (Knolle et al., 2019), and single syllables (Ott & Jäncke, 2013), the current approach was the first to use more complex bisyllabic pseudowords varying in phonotactic probability and syllable stress. We observed an interaction between phonotactic probability (high vs. low) and condition (self- vs. externally triggered) on N1 amplitude, however no post-hoc t-tests were significant. P2 amplitude was modulated by syllable stress, with first syllable stress items eliciting a larger P2 compared to second syllable stress items, likely reflecting acoustic differences between the stimuli. These findings provide preliminary evidence suggesting that phonotactic regularities influence processing of self- and externally produced speech differently, which should be followed up on in future investigations.

Taken together, the results presented in this thesis suggest that Dutch-speaking adults are sensitive to variations in phonotactic probability, indicated by variations in MMN latency (Chapter 2). This sensitivity is reduced in adult dyslexic readers (Chapter 3), suggesting that sensitivity to phonotactic regularities may play an important role in successful reading development. While Dutch adults, with and

without dyslexia, are able to perceptually distinguish between different stress patterns (Chapters 2, 3 and 4), we did not observe facilitated processing of more typical stress patterns, suggesting that it may not be a particularly relevant cue in sublexical speech processing of single pseudowords in a language with variable stress patterns. Together with the time-frequency results (Chapter 3), these findings suggest that phonotactic probability may be a more important cue for speech processing of individual pseudowords, and that the phonological and temporal regularities are processed independently at the sublexical level.

2   General discussion

Sublexical regularities in the speech signal play an important role in the acquisition of language skills. During development, we acquire implicit knowledge of regularities in the co-occurrence of speech sounds (Saffran et al., 1996) and syllable stress patterns (Jusczyk et al., 1999) of our native language. These regularities are subsequently exploited to guide the parsing of the speech signal into words (Thiessen & Saffran, 2003), and continue to influence speech processing throughout the lifespan (e.g., Edwards et al., 2004; Storkel et al., 2006; Vitevitch & Luce, 1999). Reduced sensitivity to sublexical regularities is associated with difficulties in speech perception tasks that are often observed in dyslexic readers, such as atypical categorical perception (Noordenbos et al., 2012; Serniclaes et al., 2004 but see e.g., Blomert & Mitterer, 2004; Romanovska et al., 2019) or impaired lexical stress perception (Goswami et al., 2013; Jiménez-Fernández et al., 2015; Leong et al., 2011). Thus, an understanding of how these features are processed in adult typical and dyslexic readers can contribute to understanding deficits underlying developmental dyslexia.

2.1   Phonotactic regularities shape sublexical speech processing

Phonotactic regularities have been shown to shape the neural processing of the speech signal. During development, a crucial function of a sensitivity to statistical regularities of co-occurring speech sounds is to facilitate segmentation of the speech signal into words (Saffran et al., 1996). Phonotactic probability influences speech perception (e.g., Luce & Large, 2001; Vitevitch & Luce, 1999) and production (e.g., Edwards et al., 2004; Vitevitch & Luce, 1998, 2005) across the lifespan. Even in adulthood we are able to learn new regularities (Batterink, 2017; Zhang et al., 2021),

and learned distributional statistics are encoded in low-frequency neural oscillations (Di Liberto et al., 2019). Prior studies investigating neural sensitivity to phonotactic probability in a passive oddball paradigm in adults (Bonte et al., 2005; Noordenbos et al., 2013) and children (Bonte et al., 2007) has demonstrated facilitated processing of high probability items, indexed through modulations of MMN amplitude, which is reduced in dyslexic readers (Bonte et al., 2007; Noordenbos et al., 2013).

The current thesis provides additional evidence supporting previously reported observations of facilitated change detection, as measured in the MMN, of high phonotactic probability items compared to low phonotactic probability items (Chapter 2), with a reduced sensitivity in individuals with dyslexia (Chapter 3). We used a reversed oddball paradigm, where the mismatch response is calculated as the difference between *identical* stimuli presented as deviant and standard in separate conditions (Bonte et al., 2005; Eulitz & Lahiri, 2004). This approach can isolate the effect of perceptual/auditory change detection beyond just acoustic stimulus differences. Prior research using this approach has reported this facilitated processing in the context of an amplitude modulation, with high probability items eliciting a larger mismatch negativity compared to low probability items (Bonte et al., 2005, 2007; Noordenbos et al., 2013). However, latency modulations have been observed in other paradigms that investigated the ERP correlates of phonotactic probability (Hunter, 2013; Pylkkänen et al., 2002; Stockall et al., 2004). The MMN modulation was previously interpreted as an indicator of Hebbian learning, where high probability items have established more stable auditory memory traces, which would elicit larger MMN amplitudes (Bonte et al., 2005). The current latency modulation may still be interpreted in the same context, with auditory memory traces for high probability deviants being accessed more readily. Therefore, faster neural processing of high probability items may contribute to the perceptual advantage reported in behavioral studies (Luce & Large, 2001; Vitevitch & Luce, 1999). The change from an amplitude modulation (Bonte et al., 2005; Noordenbos et al., 2013) to a latency modulation may be related to the different context in which the stimuli were presented, where prior MMN studies manipulated only phonotactic probability, but the current experiments varied both phonotactic probability and syllable stress simultaneously.

While there is ample electrophysiological evidence supporting the role of phonotactic regularities in speech perception, we know of no studies investigating their electrophysiological correlates in speech production. Functional MRI evidence suggests that distributional statistics including phonotactic probability, syllable frequencies and mutual information (quantification of statistical dependence between two syllables based on their joint probability), are represented across the speech network, including auditory as well as motor regions (Papoutsi et al., 2009; Tremblay et al., 2016). Here, reduced BOLD signal for items with higher frequency of occurrence supports psycholinguistic models proposing a "mental syllabary" storing motor plans of frequently occurring structures (Levelt, 1999; Levelt & Wheeldon, 1994; Schiller et al., 1996). This stands in contrast to our interpretation of the MMN, where greater amplitudes are interpreted as more robust memory representations for high probability items (Bonte et al., 2005; Noordenbos et al., 2013), however these differences are likely explained by differences in task demands (overt speech production vs. passive perception). The sparse neural evidence surrounding the role of sublexical regularities in speech production may be related to methodological issues in studying speech production. Engaging articulators can lead to motor artifacts in the EEG signal, and utterances containing atypical phonotactic and rhythmic structures may be associated with increased variability in the speech signal and are more prone to speech errors (Heisler & Goffman, 2016; Munson, 2001; Sasisekaran et al., 2010). These issues can complicate investigations using conventional analysis techniques. Some studies investigating motor predictions involved in speech production have applied a paradigm comparing speech stimuli that are self-triggered via button press or externally generated (Knolle et al., 2019; Ott & Jäncke, 2013; Pinheiro et al., 2018).

In Chapter 4 of the dissertation, we provide preliminary evidence suggesting distinct roles of phonotactic predictions for speech that is self-triggered via button press, or externally generated. While this approach is certainly not directly equivalent to speech production, as it does not require compilation of motor plans for articulation, it can still provide useful insights into processing of self-triggered and externally generated speech stimuli. In this experiment, we observe an interaction between phonotactic probability (high vs. low) and condition (self-triggered vs. externally presented) on N1 amplitude. While post-hoc contrasts did not reach

significance, the presence of the interaction suggests that variations in phonotactic probability affect neural processing differently across conditions. Indeed, this effect appears to be a cross-over interaction, where directionality of the difference between high and low phonotactic probability items differs across conditions: for externally generated speech, low probability items showed a slightly greater N1 amplitude compared to high probability items, while for self-triggered speech, high probability items show a greater N1 amplitude. However, the lack of significant post-hoc tests and limitations to the design of this experiment limit the strength of the provided evidence, as is discussed in Chapter 4. These issues and possible solutions will be discussed in more detail in the section "3.3 Motor-induced suppression as a marker of predictive processing".

Taken together, this dissertation provides evidence suggesting that phonotactic regularities shape sublexical speech processing. In speech perception, high probability items show facilitated processing, as indicated by shorter latencies (Emmendorfer et al., 2020) or greater amplitudes (Bonte et al., 2005; Noordenbos et al., 2013) of the MMN. This sensitivity is reduced in dyslexic children (Bonte et al., 2007) and adults (Noordenbos et al., 2013), suggesting that sensitivity to these regularities may be relevant for successful reading development. Perceptual sensitivity to phonotactic regularities can allow the listener to anticipate upcoming speech sounds and can be exploited to bias perception toward more likely utterances, or signal deviations from expected input. The neural evidence provided here may explain behavioral differences observed in speech perception and production tasks such as nonword repetition, recognition and recall (Edwards et al., 2004; Munson, 2001; Munson, Kurtz, et al., 2005; Vitevitch et al., 1997; Vitevitch & Luce, 1998, 1999, 2005). However, further investigations are needed to draw conclusions about the neural correlates of phonotactic probability in speech production.

2.2 No evidence for facilitative role of syllable stress regularities in sublexical processing of pseudowords

Regularities in the temporal structure of speech, such as variations in stress patterns provide an important cue for segmenting the continuous speech signal into words during early language acquisition (Jusczyk et al., 1999; Thiessen & Saffran, 2003), and remain relevant later in life for resolving lexical conflict (Cutler, 2005; Cutler & Van

Donselaar, 2001). The meter of the speech signal has further been shown to influence semantic (Rothermich et al., 2012) and syntactic (Schmidt-Kassow & Kotz, 2009) processing. At the sublexical level, syllable stress may also influence speech production, with more "typical" stress patterns associated with improved performance on nonword repetition tasks (Vitevitch et al., 1997). However, it is unclear whether variations in syllable stress also influence speech perception at the sublexical level.

We present the first studies, to our knowledge, to investigate the MMN sensitivity to variations in syllable stress in Dutch speakers. We do so by applying a reverse oddball paradigm, where the MMN is calculated as the difference between deviant and standard for *identical* stimuli. This approach, as previously applied by Bonte and colleagues (2005) to investigate MMN sensitivity to phonotactic probability, isolates the effect of deviance detection beyond pure acoustic effects. Most prior studies investigating MMN sensitivity to variations in syllable stress do so by comparing different stimuli to each other (e.g. SylS1 deviant - SylS2 standard; Honbolygó et al., 2004; Ylinen et al., 2009), thus the observed effect may also reflect acoustic-phonetic differences in loudness, pitch, duration, and vowel quality. A further limitation of prior studies in variable stress languages such as German (Schaadt & Männel, 2019; Weber et al., 2004) is that they do not directly statistically compare the MMN elicited by first and second syllable stress deviants to each other, but simply note the presence or absence (Weber et al., 2004), or group differences between children with and without reading difficulties (Schaadt & Männel, 2019). Thus, the current approach fills an important gap by statistically comparing MMN amplitude and latency across stress patterns.

Evidence from a comparable analysis in a fixed-stress language shows an increased MMN in response to single pseudowords with an illegal stress pattern (Honbolygó & Csépe, 2013). Similar modulations have been observed in response to violations of phonotactic constraints (Steinberg et al., 2011), suggesting that this increased MMN amplitude indexes a violation response following what might be considered a linguistic error, rather than facilitated change detection of more predictable items. As lexical stress typically, but not exclusively, falls on the first syllable in bisyllabic Dutch words, we expected the MMN modulation to follow the pattern previously observed for phonotactic probability, with first syllable stress

items eliciting a larger or earlier MMN compared to second syllable stress items. Results from Chapter 2 and 3 highlight that deviants differing from the standard in terms of syllable stress elicit an MMN, indicating that both typical and dyslexic readers can perceptually distinguish between first and second syllable stress items. However, these analyses did not reveal any significant modulation of this MMN response by variations in syllable stress.

In adults, prosodic cues such as lexical stress can be used to resolve lexical conflict in words where changes in stress patterns can alter the meaning (e.g. <u>pre</u>sent - pre<u>sent</u>; Cutler, 2005; Cutler & Van Donselaar, 2001). However, whether syllable stress influences speech processing at the sublexical level is less clear. Given that the stimuli included in the current design are pseudowords, the variations in stress pattern do not elicit any change in meaning. Evidence from German suggests that infants show distinct MMN responses to first and second syllable stress (Weber et al., 2004), but this distinction is not present in typically developing children or adults (Schaadt & Männel, 2019; Weber et al., 2004). While these studies do not statistically compare the MMN elicited by different stress patterns, they nevertheless suggest that variations in stress patterns influence speech perception more heavily early on in development, where they guide the segmentation of the continuous speech signal into words (Thiessen & Saffran, 2003), but may be less important in absence of lexical access or a sentence context (i.e., in individual pseudowords) once the language has been successfully acquired. This may explain the lack of MMN sensitivity to variations in syllable stress reported in Chapters 2 and 3 of this dissertation.

Based on theories of atypical processing of the speech amplitude envelope in dyslexic readers (Goswami, 2011), we investigated whether MMN may be modulated differently for syllable stress variations in typical vs. dyslexic readers. Prior analyses on typical readers did not reveal any MMN modulation by variations in syllable stress patterns (Chapter 2), and the analysis of group differences did not yield any significant results in Chapter 3. This stands in contrast to observations by Schaadt & Männel (2019), who reported group differences in the MMN to second syllable stress deviants. Crucially, Schaadt and Männel tested children at 4-5 years of age, prior to reading instruction, and subsequently split the group based on their later reading performance at approximately 9-10 years of age. Their findings highlight that implicit sensitivity to variations in syllable stress may be a suitable marker for early detection

of reading difficulties, in line with frameworks proposing atypical rhythm perception in dyslexia (Goswami, 2011; Ladányi, Persici et al., 2020).

While the oddball experiments failed to show ERP differences between first and second syllable stress items, Chapter 4 reports a modulation of P2 amplitude by syllable stress. In this experiment, we compared processing of pseudowords differing in phonotactic probability and syllable stress, where stimulus presentation was either self-triggered via button press, or externally triggered. Here, first syllable stress items elicit a significantly larger P2 compared to second syllable stress items. As we did not observe any interaction between syllable stress and condition (self- vs. externally triggered), this effect is likely driven by acoustic differences between first and second syllable stress items: as intensity is an acoustic marker of syllable stress, this would be expected to lead to overall larger amplitudes in both the N1 and P2 components for the first syllable stress items (Adler & Adler, 1989; Crowley & Colrain, 2004; Näätänen & Picton, 1987; Picton et al., 1978). The lack of effect in the N1 is puzzling, however this may be attributed to adaptation due to repeated stimulation, which typically affects the N1 component to a greater extent than the P2 (Kenemans et al., 1989). Additionally, with the relatively long inter-stimulus intervals of around 2.4 s, the N1 response in the current paradigm could be dominated by the "unspecific" N1 subcomponent (Näätänen & Picton, 1987; SanMiguel et al., 2013), reflecting primarily an attention-triggering, orienting response rather than specific acoustic information of the signal. However, given the temporal predictability of the stimulus onset in both conditions, this subcomponent is likely suppressed. An alternative explanation for the P2 modulation might be related to categorical perception (Bidelman et al., 2013, 2020), where second syllable stress items may be subject to changes in vowel quality within the first (unstressed) syllable, which may be associated with more perceptually ambiguous vowels. However, given that this feature was not explicitly manipulated in the current design, and the fact that vowel reduction is not a dominant marker of unstressed vowels in Dutch (van Heuven & de Jonge, 2011), this may not be a likely explanation.

In summary, the findings presented in this thesis highlight that native Dutch speakers with and without dyslexia can perceptually distinguish between pseudowords with differing stress patterns, however that there is no facilitated processing of more "typical" stress patterns, or conversely, no disadvantage at

processing "atypical" stress patterns while processing single pseudowords. Perceptual sensitivity to stress cues is important across the lifespan. In infants, it contributes to word learning by supporting segmentation of the continuous speech stream into words. In adults, this feature remains important for resolving lexical conflict in languages that allow variations in stress patterns, such as Dutch. However, in absence of lexical access or a sentence context, as is the case with our pseudoword stimuli, this feature may be less relevant to speech processing. Further research is necessary to elucidate whether the acoustic-phonetic markers of lexical stress may be associated with more perceptual ambiguity in unstressed syllables, which may provide an explanation for the P2 modulations reported in Chapter 4.

2.3   Dissociable processes underlying formal and temporal prediction in speech

Both formal and temporal regularities have been shown to influence speech and language processing throughout the lifespan, and are proposed to operate via separate, yet interacting, functional networks (Arnal & Giraud, 2012; Kotz & Schwartze, 2010). However, while these features vary simultaneously in natural speech, they are often studied in isolation, and thus the precise nature in which they influence perception remains unclear. Some evidence suggests that a regular meter can facilitate semantic (Rothermich et al., 2012) and syntactic processing (Schmidt-Kassow & Kotz, 2009) in the context of sentences. However, it remains unclear whether they interact at the sublexical level during processing of individual pseudowords.

In the oddball paradigm reported in Chapters 2 and 3, we aimed to explore whether simultaneous manipulation of phonotactic probability and syllable stress would lead to an interactive effect of both features on the mismatch response. Statistical analyses did not yield any significant interaction. However, as discussed previously, our design showed an effect of phonotactic probability on MMN latency, whereas previous studies reported a modulation of MMN amplitude (Bonte et al., 2005; Noordenbos et al., 2013). One possibility for this discrepancy might be differences in stimulus types used across studies. However, the current experiment adapted the stimuli from Bonte and colleagues (adding the manipulation of syllable stress), and therefore would expect similar ERP modulations. An alternative

explanation is that simultaneous variation of phonotactic probability and stress patterns within a condition altered the processes underlying findings reported in previous oddball studies, suggesting that syllable stress variations may still influence phonotactic processing even if this does not lead to clear interactions.

We also did not observe any direct interactions between phonotactic probability and syllable stress in our motor-to-auditory paradigm reported in Chapter 4. In fact, these features elicited modulations in separate ERP components, with phonotactic probability influencing N1 amplitude, and syllable stress the P2 amplitude. While both phonotactic probability and syllable stress varied in this experiment, stimuli within a mini-block of 14 – 28 trials differed from each other in only one feature. Thus, it remains to be shown whether manipulation of both features within a mini-block might show a context effect as was indicated in the oddball paradigm. The observation of a temporal separation of the effects of phonotactic probability and syllable stress may suggest differences in the time course of neural processing of these features, however the current design does not allow us to disentangle differences in the time course of neural processing from stimulus-specific effects. The manipulation of phonotactic probability and syllable stress occur at different time points in the stimulus, and coarticulatory cues make it difficult to define the precise timing of when information relevant to these features becomes available. Thus, while the current results suggest a temporal dissociation of the processing of phonotactic probability and syllable stress, we cannot draw strong conclusions about the relative time course of processing.

In Chapter 3 of this dissertation, we show evidence for dissociable oscillatory mechanisms underlying phonological and rhythmic processing of speech. Here, we aimed to characterize the oscillatory correlates of the mismatch response to formal and temporal deviants. Formal deviants, which differed from the standard in phonotactic probability, elicited increased delta/theta ITC relative to standards. This is consistent with prior reports of the oscillatory correlate of the MMN (Fuentemilla et al., 2008; Hsiao et al., 2009), and is typically associated with sensory memory processes related to auditory change detection. Temporal deviants on the other hand, were associated with a *decrease* in delta/theta ITC. When manipulated in isolation in simple tone stimuli, the acoustic markers of lexical stress (duration, pitch, intensity) are all associated with the typical increase in theta ITC (de la Salle et al., 2019;

Fuentemilla et al., 2008; Halliday et al., 2014). Thus, the current results suggest that deviants differing from the standard in stress pattern are processed via a different mechanism than the sensory memory processes typically underlying auditory change detection. The change in temporal structure of the temporal deviant, following a sequence of rhythmically repetitive stimuli, may disrupt the phase-locking of oscillators to the stimulus.

Taken together, the evidence presented in this thesis suggests that phonotactic probability and syllable stress variations in single pseudowords are processed via independent mechanisms. Support for this comes from the lack of interactive effects of these features on ERP measures (Chapters 2, 3 and 4), temporally dissociable modulations of auditory evoked potentials (Chapter 4), and distinct oscillatory modulations in response to formal and temporal deviants (Chapter 3). Sensitivity to formal and temporal regularities of speech is acquired at different developmental timescales. While sensitivity to rhythmic properties of the native language can already be observed in the early days after birth (Nazzi et al., 1998), sensitivity to phonological regularities, including phonotactics, does not emerge until ~6-8 months of age (Jusczyk et al., 1999; Mattys & Jusczyk, 2001; Maye et al., 2002; Thiessen & Saffran, 2003). And while neural sensitivity to phonotactic regularities persists into adulthood (Bonte et al., 2005; Emmendorfer et al., 2020; Noordenbos et al., 2013), evidence from paradigms using single pseudowords suggests this may not always be the case for rhythmic regularities in languages with variable stress patterns (Emmendorfer et al., 2020; Weber et al., 2004). These different developmental trajectories may already point towards independent neural mechanisms. Indeed, prior models of speech processing have proposed distinct functional and oscillatory networks engaged in processing the formal and temporal structure of the speech signal. While there is evidence of regularities in the speech rhythm facilitating processing of the formal structure of speech at the semantic and syntactic level (Rothermich et al., 2012; Schmidt-Kassow & Kotz, 2009), the current results suggest that these interactive effects might not yet occur at during sublexical processing of single pseudowords. Speech processing models propose that temporal information and memory representations are first integrated in the dorsolateral prefrontal cortex (DLPFC) to facilitate perception (Kotz & Schwartze, 2010, 2016). Given the functional role of the DLPFC in language processing has been mostly attributed to higher-level

processing such as pragmatics, sentence processing, or comprehension of non-literal language (for review, see Hertrich et al., 2021), this is consistent with the lack of clear interactive effects at the sublexical level.

3   Limitations and outlook

The findings reported in this dissertation suggest that regularities in the phonological and temporal structure of speech contribute to speech processing via different processes. We observe facilitated processing of items with high phonotactic probability (Chapter 2), which is absent in adult dyslexic readers (Chapter 3). On the other hand, regularities in syllable stress do not facilitate speech processing at the sublexical level in typical (Chapter 2) and dyslexic readers (Chapter 3). Interestingly, these ERP responses appear to be driven by distinct oscillatory processes for phonotactic probability and syllable stress (Chapter 3). Preliminary evidence further suggests that phonotactic regularities may influence speech processing differently depending on if it is self-triggered or externally generated, while variations in syllable stress led to ERP modulations that seem to be primarily driven by acoustic differences (Chapter 4). The current results provide useful insights into processing of sublexical regularities in speech, but also opens the door to new questions to be addressed. In the following section, three key issues will be discussed that should be taken into consideration for future approaches.

3.1   Adult dyslexic readers

In Chapter 3, we tested adult dyslexic and typical readers in a passive oddball paradigm. While post-hoc t-tests revealed facilitated processing of high phonotactic probability deviants in typical but not dyslexic readers, this was not reflected in an interaction between phonotactic probability (high vs. low) and group (dyslexics vs. controls), as has previously been reported in both adults (Noordenbos et al., 2013) and children (Bonte et al., 2007) with dyslexia. Furthermore, we did not report any atypical processing of variations in syllable stress in the dyslexic group, despite prior evidence suggesting altered MMN responses in pre-reading children who later develop reading difficulties (Schaadt & Männel, 2019), and other reports of atypical stress perception in dyslexic readers (Goswami et al., 2013; Jiménez-Fernández et al., 2015; Leong et al., 2011).

While we were able to provide additional evidence for reduced sensitivity to phonotactic probability in individuals with dyslexia, we did not find any significant associations between reading abilities and EEG outcome in correlation analyses. Previous investigations have reported significant associations between ERP measures of sensitivity to phonotactic probability and reading abilities in children (Bonte et al., 2007) but not adults (Noordenbos et al., 2013). One explanation for this discrepancy might be differences in the population sampled for these studies. While studies of dyslexia in children typically include individuals who are diagnosed with reading difficulties at an early age, studies on adults are often biased towards individuals within the university community. This limitation also applies to the current sample of dyslexic readers, as the study was advertised within the university network. While this recruitment ensured comparable educational backgrounds between typical and dyslexic readers, it also results in a sample that is likely not representative of the dyslexic population as a whole. In fact, recent analyses in Finland suggest that only 16.7% of individuals with diagnosed childhood reading difficulties go on to pursue higher education, in contrast to 41.5% without such a diagnosis (Kortteinen et al., 2021). Additionally, there are reports of increased numbers of dyslexia diagnosis in Dutch high schools (Ministerie van Onderwijs, Cultuur en Wetenschap, 2019), which exceed the expected prevalence rates (Blomert, 2005). These later diagnoses typically do not undergo the same diagnostic procedure for persistent, specific dyslexia ("ernstige enkelvoudige dyslexie") diagnosed in elementary schools, and therefore likely include poor readers who did not meet the cut-off for a formal dyslexia diagnosis during primary school (Ministerie van Onderwijs, Cultuur en Wetenschap, 2019). With 5 of 17 dyslexic readers being diagnosed at ages 14 or older, and only 7 of dyslexic readers reporting receiving any kind of intervention, it is likely that some participants may fall into this group. Future investigations should address this issue by recruiting both typical and dyslexic readers from a broader population in order to sample more broadly across educational and socioeconomic backgrounds, and other factors that may contribute to reading abilities in adulthood. Additionally, recent developments suggest moving away from a categorical dyslexic vs. non-dyslexic approach, and to instead study reading abilities along a continuum (e.g., Peters & Ansari, 2019), which would also be facilitated through broader sampling of the population.

Developmental dyslexia can have a substantial impact on the individual later in life as it is associated with poor outcomes in a range of domains including academic and professional success as well as emotional wellbeing (Livingston et al., 2018). Thus, early diagnosis and intervention is crucial for improving later life outcomes. Schaadt and Männel (2019) showed evidence suggesting atypical processing of syllable stress in pre-readers who later exhibit reading difficulties. Additionally, atypical processing of phonotactic probability has been observed in children diagnosed with dyslexia (Bonte et al., 2007). Thus, applying the current paradigm to pre-reading children may be useful to better characterize differences in the developmental trajectory of neural sensitivity to these features. Such approaches could provide insights into potential early markers of risk factors for reading difficulties that are independent from explicit task demands.

3.2   Phonotactic probability as a continuum

The aim of the experiments summarized in this thesis was to investigate the electrophysiological correlates of predictive processing of speech, with a focus on phonological and temporal regularities at the sublexical level. In the present experiments, phonotactic probability was studied in a binary classification of "high" and "low" phonotactic probability with stimuli adapted from Bonte and colleagues (2005). In Chapters 2 and 3, stimuli consisted of the pseudowords *notsal* and *notkal* (Bonte et al., 2005; Experiment 1), for high vs. low phonotactic probability, respectively, while Chapter 4 replaced the low phonotactic probability item with *notfal* to ensure closer acoustic similarity between stimuli (Bonte et al., 2005; Experiment 2). While such a categorical approach to phonotactic probability can provide first insights into neural sensitivity to these features to thoroughly understand predictive processing at the level of phonotactics, future investigations should consider systematically varying phonotactic probability along a continuum, to assess how this influences neural responses. If high phonotactic probability is associated with facilitated processing, the ERP measure (e.g., MMN latency) should correlate with phonotactic probability.

The precise way predictions are used to facilitate perception is a topic that has garnered significant debate. A recent framework by Press and colleagues (2020) highlights the "paradox" that exists in the field of predictive processing, where two

types of prediction mechanisms are proposed. Cancellation theories propose that the neural response to expected input is suppressed, so that novel or unexpected events are highlighted (e.g., Blakemore et al., 2000). In contrast, Bayesian theories posit that perception is biased toward expected events, resulting in perceptual and neural gain for predictable sensory input (e.g., de Lange et al., 2018). Press and colleagues attempted to resolve this paradox with an opposing process model, where it is proposed that the nature of the response might vary depending on the amount of surprise generated by the input: low surprise results in perceptual bias toward expected input, while high surprise signals novel input that may require an updating of the model of the world (Press et al., 2020).

This framework is particularly interesting considering the results presented in this thesis as well as prior investigations of the role of sublexical regularities in speech processing. Behavioral evidence suggests that phonotactic probability influences sublexical speech processing throughout the lifespan, with a high phonotactic probability advantage for non-word repetition, recognition and recall (e.g., Edwards et al., 2004; Vitevitch et al., 1997; Vitevitch & Luce, 1998). These results stand in contrast to reports of novel word learning, where an advantage for low probability items has been observed in adults and children when controlling for neighborhood density (Storkel, 2009; Storkel et al., 2006; Storkel & Lee, 2011). An explanation for this discrepancy is that low phonotactic probability items are more readily recognized as novel words that must be learned, i.e., that warrant updating of the model of the world. Current EEG results suggest facilitated processing of high probability items (Bonte et al., 2005; Emmendorfer et al., 2020; Noordenbos et al., 2013), while this pattern is reversed for phonotactic violations (Steinberg et al., 2011). Sampling phonotactic probability along a broad continuum in future EEG studies would be a useful approach to test how the framework put forward by Press and colleagues may apply to predictive processing of speech. Following this framework, one might observe a "reversal" of the effect of phonotactic probability for structures that are legal, but on the lower end of the range of phonotactic probability. Where an acoustically similar legal structure exists, this may also result in an initial perceptual bias toward more predictable structures before the "error" response to signal that the model of the world must be updated (Press & Yon, 2019; Yon & Press, 2017).

3.3　Motor-induced suppression as a marker of predictive processing

In Chapter 4, we aimed to investigate how variations in phonological and temporal regularities in speech influence processing of self- vs. externally triggered speech. We set out to do so using motor-induced suppression (MIS) as a measure of prediction accuracy. MIS describes the reduced sensory and neural response for self-triggered sensations compared to identical sensory input that is externally generated. However, a critical flaw in the current design prevented us from testing our initial hypotheses surrounding motor-induced suppression. The visual cue included in all conditions resulted in an undesired positive deflection in the pre-stimulus interval. While this positivity was effectively removed from the self-induced condition via the MA-MO subtraction, it remained present in the auditory-only condition, leading to differences between conditions already preceding stimulus onset. Due to this confound, the results from Chapter 4 must be interpreted with caution. However, considering the limitations of the design and the observations made, we can inform the design and selection of paradigms and stimuli to further validate the observed patterns.

First, we would consider adaptations to the paradigm. Some suggestions for adapting the current paradigm to allow for a comparison of AO and MAC conditions include a visual control condition, as well as jittering the timing of the interval between cue and stimulus presentation to reduce the influence of the cue on the ERP components. An additional adaptation to the current design would be in the trial duration, as the long ISIs as used here may only allow investigating rather *unspecific* predictions (i.e., anticipating that *a* stimulus will be presented, but not *which* specific stimulus; SanMiguel et al., 2013). While we did observe an interaction between phonotactic probability and condition, the post-hoc tests failed to reach significance. Thus, it is possible that an adaptation to the ISI would give rise to stronger evidence supporting an effect of phonotactic probability. Some studies have shown stimulus specific effects on the N1 suppression using long ISIs, however these approaches often manipulated local stimulus regularities, where the probability of the different stimuli occurring within the experiment differed, similarly to an oddball paradigm (e.g., Knolle et al., 2019). Thus, these may represent an MMN-like effect corresponding to perceptual deviance detection rather than stimulus-specific differences in suppression. In experiments with overt speech production (e.g., Niziolek et al., 2013; Sitek et al., 2013), stimulus-specific differences in suppression effects may reflect

more finely tuned motor predictions and monitoring processes compared to studies where stimuli are triggered by button-press.

While the current experiment was designed specifically to investigate modulations of motor-induced suppression, the overall aim of this research was not to understand this suppression effect per se, but rather to understand how regularities of certain features (here: phonotactic probability and syllable stress) contribute to speech production. Thus, the suppression effect observed when comparing self- and externally generated stimuli is simply one measure that can be used to address such questions related to motor predictions. Understanding how such regularities modulate processing of self-triggered stimuli can also be done without a direct comparison to passive perception (Korka et al., 2019), e.g., by varying predictability based on local stimulus regularity (i.e., as in an oddball paradigm) or action intention (whether presented stimulus corresponds to intended/expected stimulus), which would eliminate the need for the auditory-only condition.

In addition to these adaptations to the experimental paradigm, we can consider adaptations to the stimuli used. The stimuli consisted of bisyllabic Dutch pseudowords varying in phonotactic probability and syllable stress, where all variations are legal constructions in the Dutch language. The manipulation of phonotactic probability occurs at the syllable boundary, while the syllable stress contrast is already apparent from the stimulus onset. Thus, the selection of the stimuli may have been unsuitable for testing the effects of phonotactic probability on the N1 component, as the manipulation occurs later (though coarticulatory cues may already be present within the first syllable). Future designs should consider manipulating this feature already in the first syllable, to ensure that the timing of this manipulation allows it to be integrated in processes underlying N1/P2 generation. Additionally, varying the timing of the manipulation of phonotactic probability may also allow us to answer questions about the relative time course of processing phonotactic probability and syllable stress.

Despite the limitations to the design, the presence of this cue and the resulting ERP responses in the current experiment draw attention to more general limitations of such paradigms that warrant further investigations. While we cannot draw strong conclusions about a direct comparison of auditory-only and motor-auditory conditions in the current data, it is striking that we do not observe any significant

suppression of the N1 or P2 components, despite this effect being consistently reported in prior studies applying this paradigm (e.g., Bäss et al., 2008; Knolle et al., 2012, 2013a, 2013b, 2019; Ott & Jäncke, 2013; Pinheiro et al., 2018). A crucial difference between our design and the bulk of prior research using this paradigm is the inclusion of the visual cue. Critically, while the interval between cue and stimulus presentation varies in the MA condition due to variability in response time, this interval was fixed at 500 ms for the AO condition. This results in a fixed temporal association between cue and stimulus onset, allowing participants to predict the timing of the stimulus presentation equally in both conditions. Indeed, temporal predictability is rarely controlled for in suppression paradigms (Hughes et al., 2013), and may account for a portion of the classically observed suppression effect (Sowman et al., 2012). However, some studies have also reported preserved suppression effects when temporal predictability is manipulated (Bäss et al., 2008; Lange, 2011; Pinheiro et al., 2019). Some manipulated temporal predictability only in the MA condition by introducing variable delays between button-press and sound presentation (Bäss et al., 2008; Pinheiro et al., 2019), while Lange (2011) added a similar manipulation to the AO condition, varying the timing between visual cue and sound presentation. These designs therefore differ from that of Chapter 4 in that they introduce a temporal delay between button-press and sound presentation.

A recent study (Harrison et al., 2021) manipulated both temporal predictability in the AO condition, and temporal control in the MA condition by means of visual cueing, showing that MIS is preserved when both features are controlled for. The design was similar to ours in that button-press and stimulus presentation were preceded by a visual cue in the *Cued Self-generation* and *Cued Listening* conditions. However, the nature of the cue was different in that participants were instructed to synchronize their button-press with the moment in time when a vertical line moving from right to left intersects with the fixation line at the center of the screen, whereas our participants press the button in response to a visual cue, of which the timing is unpredictable. Our design thus most closely mirrors that of Sowman and colleagues (2012), who reported no differences in N1 and P2 amplitudes between cued AO and MA conditions (in Sowman et al.: *CT* and *CMT* conditions, respectively), while uncued AO conditions (in Sowman et al.: *TO* condition) showed increased amplitudes in both components. Taken together, these findings suggest that temporal predictability and

temporal control influence MIS, but the nature of these effects may depend heavily on experimental design parameters.

4   Conclusion

In three empirical chapters, this dissertation investigated electrophysiological correlates of phonological and temporal regularities in speech processing in adult typical and dyslexic readers. We provide evidence supporting previous findings (Bonte et al., 2005, 2007; Noordenbos et al., 2013) of a facilitative effect of high phonotactic probability on speech processing in a passive oddball paradigm, indexed by shorter MMN latencies for high phonotactic probability deviants compared to low phonotactic probability deviants (Chapter 2). In line with previous observations (Bonte et al., 2007; Noordenbos et al., 2013), this sensitivity was reduced in dyslexic readers (Chapter 3). In contrast, no such sensitivity was observed for variations in syllable stress, suggesting that this cue may be less relevant in processing individual pseudowords in adults (Chapters 2, 3). Time-frequency results suggest distinct processes underlying change detection for formal and temporal regularities (Chapter 3). While we did not observe any explicit interactions between phonotactic probability and syllable stress, the deviation in MMN modulation pattern from an amplitude modulation (Bonte et al., 2005, 2007; Noordenbos et al., 2013) to a latency modulation (Chapter 2) suggests that simultaneous manipulation of phonotactic probability and syllable stress my alter processing compared to when only phonotactic probability is varied. In Chapter 4, we presented preliminary evidence suggesting that regularities in the formal structure of speech are exploited differently for perception and production. Given the limitations to the design and lack of significant post-hoc tests, follow-up experiments are necessary to validate the outcome of this experiment. Future directions can build on the research presented here by testing pre-reading children to understand whether ERP markers of phonotactic probability and syllable stress perception may be suitable for detecting a risk of developing reading difficulties (see e.g., Schaadt & Männel, 2019). Additionally, sampling phonotactic probability on a continuum may allow future research to better characterize the predictive mechanisms underlying sublexical speech processing.

References

Aaltonen, O., Niemi, P., Nyrke, T., & Tuhkanen, M. (1987). Event-related brain potentials and the perception of a phonetic continuum. *Biological Psychology*, *24*(3), 197–207.

Adler, G., & Adler, J. (1989). Influence of Stimulus Intensity on AEP Components in the 80- to 200-Millisecond Latency Range. *International Journal of Audiology*, *28*(6), 316–324. https://doi.org/10.3109/00206098909081638

Aitchison, L., & Lengyel, M. (2017). With or without you: Predictive coding and Bayesian inference in the brain. *Current Opinion in Neurobiology*, *46*, 219–227. https://doi.org/10.1016/j.conb.2017.08.010

Aliu, S. O., Houde, J. F., & Nagarajan, S. S. (2009). Motor-induced suppression of the auditory cortex. *Journal of Cognitive Neuroscience*, *21*(4), 791–802. https://doi.org/10.1162/jocn.2009.21055

Anastasiou, D., & Protopapas, A. (2015). Difficulties in Lexical Stress Versus Difficulties in Segmental Phonology Among Adolescents With Dyslexia. *Scientific Studies of Reading*, *19*(1), 31–50. https://doi.org/10.1080/10888438.2014.934452

Arnal, L. H., Doelling, K. B., & Poeppel, D. (2015). Delta–Beta Coupled Oscillations Underlie Temporal Prediction Accuracy. *Cerebral Cortex*, *25*(9), 3077–3085. https://doi.org/10.1093/cercor/bhu103

Arnal, L. H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*(7), 390–398. https://doi.org/10.1016/j.tics.2012.05.003

Auer Jr, E. T., & Luce, P. A. (2005). 25 Probabilistic Phonotactics in Spoken Word Recognition. *The Handbook of Speech Perception*, 610.

Auksztulewicz, R., & Friston, K. (2016). Repetition suppression and its contextual determinants in predictive coding. *Cortex*, *80*, 125–140. https://doi.org/10.1016/j.cortex.2015.11.024

Baayen, R., Piepenbrock, R., & Gulikers, L. (2001). *WebCelex*.

Baldeweg, T. (2006). Repetition effects to sounds: Evidence for predictive coding in the auditory system. *Trends in Cognitive Sciences*, *10*(3), 93–94. https://doi.org/10.1016/j.tics.2006.01.010

Barry, J. G., Harbodt, S., Cantiani, C., Sabisch, B., & Zobay, O. (2012). Sensitivity to Lexical Stress in Dyslexia: A Case of Cognitive not Perceptual Stress: Stress Perception in Dyslexia. *Dyslexia*, *18*(3), 139–165. https://doi.org/10.1002/dys.1440

Bäss, P., Jacobsen, T., & Schröger, E. (2008). Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: Evidence for internal forward models with dynamic stimulation. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *70*(2), 137–143. https://doi.org/10.1016/j.ijpsycho.2008.06.005

Bastiaansen, M. C. M., Linden, M. van der, Keurs, M. ter, Dijkstra, T., & Hagoort, P. (2005). Theta Responses Are Involved in Lexical—Semantic Retrieval during Language Processing. *Journal of Cognitive Neuroscience*, *17*(3), 530–541. https://doi.org/10.1162/0898929053279469

Bastiaansen, M., & Hagoort, P. (2015). Frequency-based Segregation of Syntactic and Semantic Unification during Online Sentence Level Language Comprehension. *Journal of Cognitive Neuroscience*, *27*(11), 2095–2107. https://doi.org/10.1162/jocn_a_00829

Bastiaansen, M., Magyari, L., & Hagoort, P. (2010). Syntactic Unification Operations Are Reflected in Oscillatory Dynamics during On-line Sentence Comprehension. *Journal of Cognitive Neuroscience*, *22*(7), 1333–1347. https://doi.org/10.1162/jocn.2009.21283

Batterink, L. J. (2017). Rapid Statistical Learning Supporting Word Extraction From Continuous Speech. *Psychological Science*, *28*(7), 921–928. https://doi.org/10.1177/0956797617698226

Behroozmand, R., & Larson, C. R. (2011). Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. *BMC Neuroscience*, *12*, 54. https://doi.org/10.1186/1471-2202-12-54

Bian, Y., & Dell, G. S. (2020). Novel stress phonotactics are learnable by English speakers: Novel tone phonotactics are not. *Memory & Cognition*, *48*(2), 176–187. https://doi.org/10.3758/s13421-019-01000-9

Bidelman, G. M., Bush, L. C., & Boudreaux, A. M. (2020). Effects of Noise on the Behavioral and Neural Categorization of Speech. *Frontiers in Neuroscience*, *14*, 153. https://doi.org/10.3389/fnins.2020.00153

Bidelman, G. M., Moreno, S., & Alain, C. (2013). Tracing the emergence of categorical speech perception in the human auditory system. *NeuroImage*, *79*, 201–212. https://doi.org/10.1016/j.neuroimage.2013.04.093

Bishop, D. V. M., Hardiman, M. J., & Barry, J. G. (2011). Is auditory discrimination mature by middle childhood? A study using time-frequency analysis of mismatch responses from 7 years to adulthood. *Developmental Science*, *14*(2), 402–416. https://doi.org/10.1111/j.1467-7687.2010.00990.x

Blakemore, S. J., Wolpert, D., & Frith, C. (2000). Why can't you tickle yourself? *Neuroreport*, *11*(11), R11-16. https://doi.org/10.1097/00001756-200008030-00002

Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2018). Effect of variance ratio on ANOVA robustness: Might 1.5 be the limit? *Behavior Research Methods*, *50*(3), 937–962. https://doi.org/10.3758/s13428-017-0918-2

Bles, M., Alink, A., & Jansma, B. M. (2007). Neural aspects of cohort-size reduction during visual gating. *Brain Research*, *1150*, 143–154. https://doi.org/10.1016/j.brainres.2007.02.071

Blomert, L. (2005). *Dyslexie in Nederland*. Uitgeverij Nieuwezijds.

Blomert, L. (2011). The neural signature of orthographic–phonological binding in successful and failing reading development. *Neuroimage*, *57*(3), 695–703.

Blomert, L., & Mitterer, H. (2004). The fragile nature of the speech-perception deficit in dyslexia: Natural vs. synthetic speech. *Brain and Language*, *89*(1), 21–26. https://doi.org/10.1016/S0093-934X(03)00305-5

Boersma, P., & Weenink, D. (2013). *Praat: Doing phonetics by computer. Version 5.3. 55. Http*.

Bonte, M. L., Mitterer, H., Zellagui, N., Poelmans, H., & Blomert, L. (2005). Auditory cortical tuning to statistical regularities in phonology. *Clinical Neurophysiology*, *116*(12), 2765–2774. https://doi.org/10.1016/j.clinph.2005.08.012

Bonte, M. L., Poelmans, H., & Blomert, L. (2007). Deviant neurophysiological responses to phonological regularities in speech in dyslexic children. *Neuropsychologia*, *45*(7), 1427–1437.

Breska, A., & Deouell, L. Y. (2017). Neural mechanisms of rhythm-based temporal prediction: Delta phase-locking reflects temporal predictability but not rhythmic entrainment. *PLOS Biology*, *15*(2), e2001665. https://doi.org/10.1371/journal.pbio.2001665

Bressler, S. L., & Richter, C. G. (2015). Interareal oscillatory synchronization in top-down neocortical processing. *Current Opinion in Neurobiology*, *31*, 62–66. https://doi.org/10.1016/j.conb.2014.08.010

Brodski, A., Paasch, G.-F., Helbling, S., & Wibral, M. (2015). The Faces of Predictive Coding. *Journal of Neuroscience*, *35*(24), 8997–9006. https://doi.org/10.1523/JNEUROSCI.1529-14.2015

Brus, B. T., & Voeten, M. J. M. (1973). *Een-minuut-test: Vorm A en B; schoolvorderingentest voor de technische leesvaardigheid, bestemd voor het tweede tot en met het zesde leerjaar van het basisonderwijs; verantwoording en handleiding:*

Cantiani, C., Ortiz-Mantilla, S., Riva, V., Piazza, C., Bettoni, R., Musacchia, G., Molteni, M., Marino, C., & Benasich, A. A. (2019). Reduced left-lateralized pattern of event-related EEG oscillations in infants at familial risk for language and learning impairment. *NeuroImage. Clinical*, *22*, 101778. https://doi.org/10.1016/j.nicl.2019.101778

Carpenter, A. L., & Shahin, A. J. (2013). Development of the N1–P2 auditory evoked response to amplitude rise time and rate of formant transition of speech sounds. *Neuroscience Letters*, *544*, 56–61. https://doi.org/10.1016/j.neulet.2013.03.041

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience*, *1*(5), 351–353. https://doi.org/10.1038/1561

Christoffels, I. K., van de Ven, V., Waldorp, L. J., Formisano, E., & Schiller, N. O. (2011). The sensory consequences of speaking: Parametric neural cancellation during speech in auditory cortex. *PloS One*, *6*(5), e18307. https://doi.org/10.1371/journal.pone.0018307

Cohen, M. X. (2014). *Analyzing neural time series data: Theory and practice*. The MIT Press.

Conde, T., Gonçalves, Ó. F., & Pinheiro, A. P. (2018). Stimulus complexity matters when you hear your own voice: Attention effects on self-generated voice processing. *International Journal of Psychophysiology*, *133*, 66–78. https://doi.org/10.1016/j.ijpsycho.2018.08.007

Cone-Wesson, B., & Wunderlich, J. (2003). Auditory evoked potentials from the cortex: Audiology applications: *Current Opinion in Otolaryngology & Head and Neck Surgery*, *11*(5), 372–377. https://doi.org/10.1097/00020840-200310000-00011

Costa-Faidella, J., Baldeweg, T., Grimm, S., & Escera, C. (2011). Interactions between "What" and "When" in the Auditory System: Temporal Predictability Enhances Repetition Suppression. *Journal of Neuroscience*, *31*(50), 18590–18597. https://doi.org/10.1523/JNEUROSCI.2599-11.2011

Cramer, A. O. J., van Ravenzwaaij, D., Matzke, D., Steingroever, H., Wetzels, R., Grasman, R. P. P. P., Waldorp, L. J., & Wagenmakers, E.-J. (2016). Hidden multiplicity in exploratory multiway ANOVA: Prevalence and remedies. *Psychonomic Bulletin & Review*, *23*(2), 640–647. https://doi.org/10.3758/s13423-015-0913-5

Cravo, A. M., Rohenkohl, G., Wyart, V., & Nobre, A. C. (2011). Endogenous modulation of low frequency oscillations by temporal expectations. *Journal of Neurophysiology*, *106*(6), 2964–2972. https://doi.org/10.1152/jn.00157.2011

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *115*(4), 732–744. https://doi.org/10.1016/j.clinph.2003.11.021

Cummins, F., & Port, R. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics*, *26*(2), 145–171. https://doi.org/10.1006/jpho.1998.0070

Cutler, A. (2005). *Lexical stress*.

Cutler, A., & Van Donselaar, W. (2001). Voornaam is not (really) a Homophone: Lexical Prosody and Lexical Access in Dutch. *Language and Speech*, *44*(2), 171–195. https://doi.org/10.1177/00238309010440020301

Dalla Bella, S., Farrugia, N., Benoit, C.-E., Begel, V., Verga, L., Harding, E., & Kotz, S. A. (2017). BAASTA: Battery for the Assessment of Auditory Sensorimotor and Timing Abilities. *Behavior Research Methods*, *49*(3), 1128–1145. https://doi.org/10.3758/s13428-016-0773-6

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, *11*(1), 51–62. https://doi.org/10.1016/S0095-4470(19)30776-4

de la Salle, S., Shah, D., Choueiry, J., Bowers, H., McIntosh, J., Ilivitsky, V., & Knott, V. (2019). NMDA Receptor Antagonist Effects on Speech-Related Mismatch Negativity and Its Underlying Oscillatory and Source Activity in Healthy Humans. *Frontiers in Pharmacology*, *10*, 455. https://doi.org/10.3389/fphar.2019.00455

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How Do Expectations Shape Perception? *Trends in Cognitive Sciences*, *22*(9), 764–779. https://doi.org/10.1016/j.tics.2018.06.002

Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults: *NeuroReport*, *8*(4), 919–924. https://doi.org/10.1097/00001756-199703030-00021

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*(8), 1117–1121. https://doi.org/10.1038/nn1504

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Delorme, A., Sejnowski, T., & Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *Neuroimage*, *34*(4), 1443–1449.

Di Liberto, G. M., Wong, D., Melnik, G. A., & de Cheveigné, A. (2019). Low-frequency cortical responses to natural speech reflect probabilistic phonotactics. *NeuroImage*, *196*, 237–247. https://doi.org/10.1016/j.neuroimage.2019.04.037

Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music. *Neuroscience & Biobehavioral Reviews*, *81*, 181–187. https://doi.org/10.1016/j.neubiorev.2017.02.011

Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research: JSLHR*, *47*(2), 421–436. https://doi.org/10.1044/1092-4388(2004/034)

Ellis, R. J., & Jones, M. R. (2010). Rhythmic context modulates foreperiod effects. *Attention, Perception & Psychophysics*, *72*(8), 2274–2288. https://doi.org/10.3758/bf03196701

Emmendorfer, A. K., Correia, J. M., Jansma, B. M., Kotz, S. A., & Bonte, M. (2020). ERP mismatch response to phonological and temporal regularities in speech. *Scientific Reports*, *10*(1), 9917. https://doi.org/10.1038/s41598-020-66824-x

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top–down processing. *Nature Reviews Neuroscience*, *2*(10), 704–716. https://doi.org/10.1038/35094565

Eulitz, C., & Lahiri, A. (2004). Neurobiological Evidence for Abstract Phonological Representations in the Mental Lexicon during Speech Recognition. *Journal of Cognitive Neuroscience*, *16*(4), 577–583. https://doi.org/10.1162/089892904323057308

Fitzgerald, K., & Todd, J. (2020). Making Sense of Mismatch Negativity. *Frontiers in Psychiatry*, *11*, 468. https://doi.org/10.3389/fpsyt.2020.00468

Freunberger, D., & Roehm, D. (2016). Semantic prediction in language comprehension: Evidence from brain potentials. *Language, Cognition and Neuroscience*, *31*(9), 1193–1205. https://doi.org/10.1080/23273798.2016.1205202

Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9*(10), 474–480. https://doi.org/10.1016/j.tics.2005.08.011

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, *13*(7), 293–301. https://doi.org/10.1016/j.tics.2009.04.005

Friston, K. (2012). Prediction, perception and agency. *International Journal of Psychophysiology*, *83*(2), 248–252. https://doi.org/10.1016/j.ijpsycho.2011.11.014

Fuentemilla, L., Marco-Pallarés, J., Münte, T. F., & Grau, C. (2008). Theta EEG oscillatory activity and auditory change detection. *Brain Research*, *1220*, 93–101. https://doi.org/10.1016/j.brainres.2007.07.079

Gabay, Y., Gabay, S., Schiff, R., & Henik, A. (2020). Visual and Auditory Interference Control of Attention in Developmental Dyslexia. *Journal of the International Neuropsychological Society*, *26*(4), 407–417. https://doi.org/10.1017/S135561771900122X

Gabay, Y., Thiessen, E. D., & Holt, L. L. (2015). Impaired Statistical Learning in Developmental Dyslexia. *Journal of Speech, Language, and Hearing Research*, *58*(3), 934–945. https://doi.org/10.1044/2015_JSLHR-L-14-0324

García-Larrea, L., Lukaszewicz, A.-C., & Mauguiére, F. (1992). Revisiting the oddball paradigm. Non-target vs neutral stimuli and the evaluation of ERP attentional effects. *Neuropsychologia*, *30*(8), 723–741. https://doi.org/10.1016/0028-3932(92)90042-K

Gauvin, H. S., & Hartsuiker, R. J. (2020). Towards a New Model of Verbal Monitoring. *Journal of Cognition*, *3*(1), 17. https://doi.org/10.5334/joc.81

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511.

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, *15*(1), 3–10.

Goswami, U., Mead, N., Fosker, T., Huss, M., Barnes, L., & Leong, V. (2013). Impaired perception of syllable stress in children with dyslexia: A longitudinal study. *Journal of Memory and Language*, *69*(1), 1–17.

Gu, C., & Bi, H.-Y. (2020). Auditory processing deficit in individuals with dyslexia: A meta-analysis of mismatch negativity. *Neuroscience & Biobehavioral Reviews*, *116*, 396–405. https://doi.org/10.1016/j.neubiorev.2020.06.032

Halliday, L. F., Barry, J. G., Hardiman, M. J., & Bishop, D. V. (2014). Late, not early mismatch responses to changes in frequency are reduced or deviant in children with dyslexia: An event-related potential study. *Journal of Neurodevelopmental Disorders*, *6*(1), 21. https://doi.org/10.1186/1866-1955-6-21

Hannemann, R., Obleser, J., & Eulitz, C. (2007). Top-down knowledge supports the retrieval of lexical information from degraded speech. *Brain Research*, *1153*, 134–143. https://doi.org/10.1016/j.brainres.2007.03.069

Harrison, A. W., Mannion, D. J., Jack, B. N., Griffiths, O., Hughes, G., & Whitford, T. J. (2021). Sensory attenuation is modulated by the contrasting effects of predictability and control. *NeuroImage*, *237*, 118103. https://doi.org/10.1016/j.neuroimage.2021.118103

Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., & Ford, J. M. (2005). Fine-tuning of auditory cortex during speech production. *Psychophysiology*, *42*(2), 180–190. https://doi.org/10.1111/j.1469-8986.2005.00272.x

Heisler, L., & Goffman, L. (2016). The Influence of Phonotactic Probability and Neighborhood Density on Children's Production of Newly Learned Words. *Language Learning and Development*, *12*(3), 338–356. https://doi.org/10.1080/15475441.2015.1117977

Hertrich, I., Dietrich, S., Blum, C., & Ackermann, H. (2021). The Role of the Dorsolateral Prefrontal Cortex for Speech and Language Processing. *Frontiers in Human Neuroscience*, *15*, 645209. https://doi.org/10.3389/fnhum.2021.645209

Heynckes, M., De Weerd, P., Valente, G., Formisano, E., & De Martino, F. (2020). Behavioral effects of rhythm, carrier frequency and temporal cueing on the perception of sound sequences. *PloS One*, *15*(6), e0234251. https://doi.org/10.1371/journal.pone.0234251

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, *13*(2), 135–145. https://doi.org/10.1038/nrn3158

Holliman, A. J., Wood, C., & Sheehy, K. (2010). The contribution of sensitivity to speech rhythm and non-speech rhythm to early reading development. *Educational Psychology*, *30*(3), 247–267. https://doi.org/10.1080/01443410903560922

Honbolygó, F., & Csépe, V. (2013). Saliency or template? ERP evidence for long-term representation of word stress. *International Journal of Psychophysiology*, *87*(2), 165–172.

Honbolygó, F., Csépe, V., & Ragó, A. (2004). Suprasegmental speech cues are automatically processed by the human brain: A mismatch negativity study. *Neuroscience Letters*, *363*(1), 84–88.

Hsiao, F.-J., Wu, Z.-A., Ho, L.-T., & Lin, Y.-Y. (2009). Theta oscillation during auditory change detection: An MEG study. *Biological Psychology*, *81*(1), 58–66. https://doi.org/10.1016/j.biopsycho.2009.01.007

Hughes, G., Desantis, A., & Waszak, F. (2013). Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychological Bulletin*, *139*(1), 133–151. https://doi.org/10.1037/a0028566

Hunter, C. R. (2013). Early effects of neighborhood density and phonotactic probability of spoken words on event-related potentials. *Brain and Language*, *127*(3), 463–474.

Jensen, O., Bonnefond, M., & VanRullen, R. (2012). An oscillatory mechanism for prioritizing salient unattended stimuli. *Trends in Cognitive Sciences*, *16*(4), 200–206. https://doi.org/10.1016/j.tics.2012.03.002

Jiménez-Fernández, G., Gutiérrez-Palma, N., & Defior, S. (2015). Impaired stress awareness in Spanish children with developmental dyslexia. *Research in Developmental Disabilities*, *37*, 152–161. https://doi.org/10.1016/j.ridd.2014.11.002

Jin, Y., Díaz, B., Colomer, M., & Sebastián-Gallés, N. (2014). Oscillation encoding of individual differences in speech perception. *PloS One*, *9*(7), e100901. https://doi.org/10.1371/journal.pone.0100901

Johnson, J. F., Belyk, M., Schwartze, M., Pinheiro, A. P., & Kotz, S. A. (2021). Expectancy changes the self-monitoring of voice identity. *European Journal of Neuroscience*, *53*(8), 2681–2695. https://doi.org/10.1111/ejn.15162

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in english-learning infants. *Cognitive Psychology*, *39*(3–4), 159–207. https://doi.org/10.1006/cogp.1999.0716

Kassambara, A. (2019). *rstatix: Pipe-Friendly Framework for Basic Statistical Tests*.

Katseff, S., Houde, J., & Johnson, K. (2012). Partial Compensation for Altered Auditory Feedback: A Tradeoff with Somatosensory Feedback? *Language and Speech*, *55*(2), 295–308. https://doi.org/10.1177/0023830911417802

Kenemans, J. L., Verbaten, M. N., Roelofs, J.-W., & Slangen, J. L. (1989). "Initial-" and "change-orienting reactions": An analysis based on visual single-trial event-related potentials. *Biological Psychology*, *28*(3), 199–226. https://doi.org/10.1016/0301-0511(89)90001-X

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, *16*(12), 606–617. https://doi.org/10.1016/j.tics.2012.10.007

Knolle, F., Schröger, E., Baess, P., & Kotz, S. A. (2012). The cerebellum generates motor-to-auditory predictions: ERP lesion evidence. *Journal of Cognitive Neuroscience*, *24*(3), 698–706. https://doi.org/10.1162/jocn_a_00167

Knolle, F., Schröger, E., & Kotz, S. A. (2013a). Prediction errors in self- and externally-generated deviants. *Biological Psychology*, *92*(2), 410–416. https://doi.org/10.1016/j.biopsycho.2012.11.017

Knolle, F., Schröger, E., & Kotz, S. A. (2013b). Cerebellar contribution to the prediction of self-initiated sounds. *Cortex*, *49*(9), 2449–2461. https://doi.org/10.1016/j.cortex.2012.12.012

Knolle, F., Schwartze, M., Schröger, E., & Kotz, S. A. (2019). Auditory Predictions and Prediction Errors in Response to Self-Initiated Vowels. *Frontiers in Neuroscience*, *13*, 1146. https://doi.org/10.3389/fnins.2019.01146

Koerner, T. K., & Zhang, Y. (2015). Effects of background noise on inter-trial phase coherence and auditory N1–P2 responses to speech stimuli. *Hearing Research*, *328*, 113–119. https://doi.org/10.1016/j.heares.2015.08.002

Koerner, T. K., Zhang, Y., Nelson, P. B., Wang, B., & Zou, H. (2017). Neural indices of phonemic discrimination and sentence-level speech intelligibility in quiet and noise: A P3 study. *Hearing Research*, *350*, 58–67. https://doi.org/10.1016/j.heares.2017.04.009

Korka, B., Schröger, E., & Widmann, A. (2019). Action Intention-based and Stimulus Regularity-based Predictions: Same or Different? *Journal of Cognitive Neuroscience*, *31*(12), 1917–1932. https://doi.org/10.1162/jocn_a_01456

Kortteinen, H., Eklund, K., Eloranta, A., & Aro, T. (2021). Cognitive and non-cognitive factors in educational and occupational outcomes—Specific to reading disability? *Dyslexia*, *27*(2), 204–223. https://doi.org/10.1002/dys.1673

Kotz, S. A. E., & Schwartze, M. (2011). Differential input of the supplementary motor area to a dedicated temporal processing network: Functional and clinical implications. *Frontiers in Integrative Neuroscience*, *5*, 86. https://doi.org/10.3389/fnint.2011.00086

Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The Evolution of Rhythm Processing. *Trends in Cognitive Sciences*, *22*(10), 896–910. https://doi.org/10.1016/j.tics.2018.08.002

Kotz, S. A., & Schwartze, M. (2010). Cortical speech processing unplugged: A timely subcortico-cortical framework. *Trends in Cognitive Sciences*, *14*(9), 392–399. https://doi.org/10.1016/j.tics.2010.06.005

Kotz, S. A., & Schwartze, M. (2016). Chapter 57—Motor-Timing and Sequencing in Speech Production: A General-Purpose Framework. In G. Hickok & S. L. Small (Eds.), *Neurobiology of Language* (pp. 717–724). Academic Press. https://doi.org/10.1016/B978-0-12-407794-2.00057-2

Kutas, M., & Hillyard, S. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205. https://doi.org/10.1126/science.7350657

Ladányi, E., Persici, V., Fiveash, A., Tillmann, B., & Gordon, R. L. (2020). Is atypical rhythm a risk factor for developmental speech and language disorders? *WIREs Cognitive Science*, *n/a*(n/a), e1528. https://doi.org/10.1002/wcs.1528

Lallier, M., Molinaro, N., Lizarazu, M., Bourguignon, M., & Carreiras, M. (2017). Amodal atypical neural oscillatory activity in dyslexia: A cross-linguistic perspective. *Clinical Psychological Science*, *5*(2), 379–401.

Lange, K. (2011). The reduced N1 to self-generated tones: An effect of temporal predictability?: N1 reduction and temporal predictability. *Psychophysiology*, *48*(8), 1088–1095. https://doi.org/10.1111/j.1469-8986.2010.01174.x

Larson, C. R., Sun, J., & Hain, T. C. (2007). Effects of simultaneous perturbations of voice pitch and loudness feedback on voice F0 and amplitude control. *The Journal of the Acoustical Society of America*, *121*(5), 2862–2872. https://doi.org/10.1121/1.2715657

Lawrance, E. L. A., Harper, N. S., Cooke, J. E., & Schnupp, J. W. H. (2014). Temporal predictability enhances auditory detection. *The Journal of the Acoustical Society of America*, *135*(6), EL357–EL363. https://doi.org/10.1121/1.4879667

Leong, V., Hämäläinen, J., Soltész, F., & Goswami, U. (2011). Rise time perception and detection of syllable stress in adults with developmental dyslexia. *Journal of Memory and Language*, *64*(1), 59–73.

Levelt, null. (1999). Models of word production. *Trends in Cognitive Sciences*, *3*(6), 223–232. https://doi.org/10.1016/s1364-6613(99)01319-4

Levelt, W. (1983). Monitoring and self-repair in speech. *Cognition*, *14*(1), 41–104. https://doi.org/10.1016/0010-0277(83)90026-4

Levelt, W. J. M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, *50*(1–3), 239–269. https://doi.org/10.1016/0010-0277(94)90030-2

Livingston, E. M., Siegel, L. S., & Ribary, U. (2018). Developmental dyslexia: Emotional impact and consequences. *Australian Journal of Learning Difficulties*, *23*(2), 107–135. https://doi.org/10.1080/19404158.2018.1479975

Luce, P. A., & Large, N. R. (2001). Phonotactics, density, and entropy in spoken word recognition. *Language and Cognitive Processes*, *16*(5–6), 565–581. https://doi.org/10.1080/01690960143000137

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't): How to get significant effects. *Psychophysiology*, *54*(1), 146–157. https://doi.org/10.1111/psyp.12639

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, *53*(1), 1–14. https://doi.org/10.1007/s11881-003-0001-9

Mair, P., & Wilcox, R. (2020). Robust statistical methods in R using the WRS2 package. *Behavior Research Methods*, *52*(2), 464–488. https://doi.org/10.3758/s13428-019-01246-w

Mangiafico, S. (2021). *Rcompanion: Functions to Support Extension Education Program Evaluation. R package* (version 2.4.1.) [Computer software]. version 2.4.1. https://CRAN.R-project.org/package=rcompanion

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Martin, C. D., Molnar, M., & Carreiras, M. (2016). The proactive bilingual brain: Using interlocutor identity to generate predictions for language processing. *Scientific Reports*, *6*(1), 26171. https://doi.org/10.1038/srep26171

Mattys, S. L., & Jusczyk, P. W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, *78*(2), 91–121. https://doi.org/10.1016/s0010-0277(00)00109-8

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

Melby-Lervåg, M., Lyster, S.-A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, *138*(2), 322–352. https://doi.org/10.1037/a0026744

Mellem, M. S., Friedman, R. B., & Medvedev, A. V. (2013). Gamma- and theta-band synchronization during semantic priming reflect local and long-range lexical–semantic networks. *Brain and Language*, *127*(3), 440–451. https://doi.org/10.1016/j.bandl.2013.09.003

Meyer, L., & Schaadt, G. (2020). Aberrant Prestimulus Oscillations in Developmental Dyslexia Support an Underlying Attention Shifting Deficit. *Cerebral Cortex Communications*, *1*(1), tgaa006. https://doi.org/10.1093/texcom/tgaa006

Ministerie van Onderwijs, Cultuur en Wetenschap. (2019). *Dyslexieverklaringen: Verschillen tussen scholen nader bekeken*. Ministerie van Onderwijs, Cultuur en Wetenschap.

Morillon, B., & Schroeder, C. E. (2015). Neuronal oscillations as a mechanistic substrate of auditory temporal prediction: Neuronal oscillations and temporal predictions. *Annals of the New York Academy of Sciences*, *1337*(1), 26–31. https://doi.org/10.1111/nyas.12629

Morillon, B., Schroeder, C. E., Wyart, V., & Arnal, L. H. (2016). Temporal Prediction in lieu of Periodic Stimulation. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(8), 2342–2347. https://doi.org/10.1523/JNEUROSCI.0836-15.2016

Mundy, I. R., & Carroll, J. M. (2012). Speech prosody and developmental dyslexia: Reduced phonological awareness in the context of intact phonological representations. *Journal of Cognitive Psychology*, *24*(5), 560–581. https://doi.org/10.1080/20445911.2012.662341

Mundy, I. R., & Carroll, J. M. (2013). Spelling–stress regularity effects are intact in developmental dyslexia. *Quarterly Journal of Experimental Psychology*, *66*(4), 816–828. https://doi.org/10.1080/17470218.2012.719530

Munson, B. (2001). Phonological pattern frequency and speech production in adults and children. *Journal of Speech, Language, and Hearing Research: JSLHR*, *44*(4), 778–792. https://doi.org/10.1044/1092-4388(2001/061)

Munson, B., Edwards, J., & Beckman, M. E. (2005). Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. *Journal of Speech, Language, and Hearing Research: JSLHR*, *48*(1), 61–78. https://doi.org/10.1044/1092-4388(2005/006)

Munson, B., Kurtz, B. A., & Windsor, J. (2005). The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research: JSLHR*, *48*(5), 1033–1047. https://doi.org/10.1044/1092-4388(2005/072)

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

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*(12), 2544–2590.

Näätänen, R., & Picton, T. (1987). The N1 Wave of the Human Electric and Magnetic Response to Sound: A Review and an Analysis of the Component Structure. *Psychophysiology*, *24*(4), 375–425. https://doi.org/10.1111/j.1469-8986.1987.tb00311.x

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology. Human Perception and Performance*, *24*(3), 756–766. https://doi.org/10.1037//0096-1523.24.3.756

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., Von Grebmer Zu Wolfsthurn, S., Bartolozzi, F., Kogan, V., Ito, A., Mézière, D., Barr, D. J., Rousselet, G. A., Ferguson, H. J., Busch-Moreno, S., Fu, X., Tuomainen, J., Kulakova, E., Husband, E. M., … Huettig, F. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *ELife*, *7*, e33468. https://doi.org/10.7554/eLife.33468

Niziolek, C. A., Nagarajan, S. S., & Houde, J. F. (2013). What does motor efference copy represent? Evidence from speech production. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(41), 16110–16116. https://doi.org/10.1523/JNEUROSCI.2137-13.2013

Nobre, A. C., & van Ede, F. (2018). Anticipated moments: Temporal structure in attention. *Nature Reviews. Neuroscience*, *19*(1), 34–48. https://doi.org/10.1038/nrn.2017.141

Noordenbos, M., Segers, E., Serniclaes, W., Mitterer, H., & Verhoeven, L. (2012). Allophonic mode of speech perception in Dutch children at risk for dyslexia: A longitudinal study. *Research in Developmental Disabilities*, *33*(5), 1469–1483.

Noordenbos, M. W., Segers, E., Mitterer, H., Serniclaes, W., & Verhoeven, L. (2013). Deviant neural processing of phonotactic probabilities in adults with dyslexia. *Neuroreport*, *24*(13), 746–750.

Nooteboom, S. G., & Quené, H. (2017). Self-monitoring for speech errors: Two-stage detection and repair with and without auditory feedback. *Journal of Memory and Language*, *95*, 19–35. https://doi.org/10.1016/j.jml.2017.01.007

Onishi, S., & Davis, H. (1968). Effects of duration and rise time of tone bursts on evoked V potentials. *The Journal of the Acoustical Society of America*, *44*(2), 582–591. https://doi.org/10.1121/1.1911124

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, *2011*, 1–9. https://doi.org/10.1155/2011/156869

Ott, C. G. M., & Jäncke, L. (2013). Processing of self-initiated speech-sounds is different in musicians. *Frontiers in Human Neuroscience*, *7*, 41. https://doi.org/10.3389/fnhum.2013.00041

Papoutsi, M., de Zwart, J. A., Jansma, J. M., Pickering, M. J., Bednar, J. A., & Horwitz, B. (2009). From Phonemes to Articulatory Codes: An fMRI Study of the Role of Broca's Area in Speech Production. *Cerebral Cortex*, *19*(9), 2156–2165. https://doi.org/10.1093/cercor/bhn239

Parviainen, T., Helenius, P., & Salmelin, R. (2005). Cortical differentiation of speech and nonspeech sounds at 100 ms: Implications for dyslexia. *Cerebral Cortex*, *15*(7), 1054–1063. https://doi.org/10.1093/cercor/bhh206

Peelle, J. E., & Sommers, M. S. (2015). Prediction and constraint in audiovisual speech perception. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *68*, 169–181. https://doi.org/10.1016/j.cortex.2015.03.006

Peña, M., & Melloni, L. (2012). Brain Oscillations during Spoken Sentence Processing. *Journal of Cognitive Neuroscience*, *24*(5), 1149–1164. https://doi.org/10.1162/jocn_a_00144

Peperkamp, S., Le Calvez, R., Nadal, J.-P., & Dupoux, E. (2006). The acquisition of allophonic rules: Statistical learning with linguistic constraints. *Cognition*, *101*(3), B31–B41. https://doi.org/10.1016/j.cognition.2005.10.006

Peters, L., & Ansari, D. (2019). Are specific learning disorders truly specific, and are they disorders? *Trends in Neuroscience and Education*, *17*, 100115. https://doi.org/10.1016/j.tine.2019.100115

Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., McGinnis, M., & Roberts, T. (2000). Auditory Cortex Accesses Phonological Categories: An MEG Mismatch Study. *Journal of Cognitive Neuroscience*, *12*(6), 1038–1055. https://doi.org/10.1162/08989290051137567

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(4), 329–347. https://doi.org/10.1017/S0140525X12001495

Picton, T. W., Woods, D. L., Baribeau-Braun, J., & Healey, T. M. (1976). Evoked potential audiometry. *The Journal of Otolaryngology*, *6*(2), 90–119.

Picton, T. W., Woods, D. L., & Proulx, G. B. (1978). Human auditory sustained potentials. II. Stimulus relationships. *Electroencephalography and Clinical Neurophysiology*, *45*(2), 198–210. https://doi.org/10.1016/0013-4694(78)90004-4

Pierrehumbert, J. B. (2003). Phonetic Diversity, Statistical Learning, and Acquisition of Phonology. *Language and Speech*, *46*(2–3), 115–154. https://doi.org/10.1177/00238309030460020501

Pike, K. L. (1962). Practical Phonetics of Rhythm Waves. *Phonetica*, *8*(1–3), 9–30. https://doi.org/10.1159/000258116

Pinheiro, A. P., Schwartze, M., Gutierrez, F., & Kotz, S. A. (2019). When temporal prediction errs: ERP responses to delayed action-feedback onset. *Neuropsychologia*, *134*, 107200. https://doi.org/10.1016/j.neuropsychologia.2019.107200

Pinheiro, A. P., Schwartze, M., & Kotz, S. A. (2018). Voice-selective prediction alterations in nonclinical voice hearers. *Scientific Reports*, *8*(1), 14717. https://doi.org/10.1038/s41598-018-32614-9

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time.' *Speech Communication*, *41*(1), 245–255. https://doi.org/10.1016/S0167-6393(02)00107-3

Poeppel, D., Yellin, E., Phillips, C., Roberts, T. P. L., Rowley, H. A., Wexler, K., & Marantz, A. (1996). Task-induced asymmetry of the auditory evoked M100 neuromagnetic field elicited by speech sounds. *Cognitive Brain Research*, *4*(4), 231–242. https://doi.org/10.1016/S0926-6410(96)00643-X

Port, R. F. (2003). Meter and speech. *Journal of Phonetics*, *31*(3), 599–611. https://doi.org/10.1016/j.wocn.2003.08.001

Press, C., Kok, P., & Yon, D. (2020). The Perceptual Prediction Paradox. *Trends in Cognitive Sciences*, *24*(1), 13–24. https://doi.org/10.1016/j.tics.2019.11.003

Press, C., & Yon, D. (2019). Perceptual Prediction: Rapidly Making Sense of a Noisy World. *Current Biology*, *29*(15), R751–R753. https://doi.org/10.1016/j.cub.2019.06.054

Pulvermüller, F., & Shtyrov, Y. (2003). Automatic processing of grammar in the human brain as revealed by the mismatch negativity. *Neuroimage*, *20*(1), 159–172.

Purdy, S. C., Kelly, A. S., & Thorne, P. R. (2001). Auditory Evoked Potentials as Measures of Plasticity in Humans. *Audiology and Neuro-Otology*, *6*(4), 211–215. https://doi.org/10.1159/000046835

Pylkkänen, L., Stringfellow, A., & Marantz, A. (2002). Neuromagnetic evidence for the timing of lexical activation: An MEG component sensitive to phonotactic probability but not to neighborhood density. *Brain and Language*, *81*(1–3), 666–678.

R Core Team. (2013). *R: A language and environment for statistical computing*.

Ragó, A., Honbolygó, F., Róna, Z., Beke, A., & Csépe, V. (2014). Effect of maturation on suprasegmental speech processing in full-and preterm infants: A mismatch negativity study. *Research in Developmental Disabilities*, *35*(1), 192–202.

Raichle, M. E. (2010). Two views of brain function. *Trends in Cognitive Sciences*, *14*(4), 180–190. https://doi.org/10.1016/j.tics.2010.01.008

Ramus, F. (2001). Talk of two theories. *Nature*, *412*(6845), 393–394.

Ramus, F. (2003). Developmental dyslexia: Specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology*, *13*(2), 212–218.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Reznik, D., Simon, S., & Mukamel, R. (2018). Predicted sensory consequences of voluntary actions modulate amplitude of preceding readiness potentials. *Neuropsychologia*, *119*, 302–307. https://doi.org/10.1016/j.neuropsychologia.2018.08.028

Rimmele, J. M., Morillon, B., Poeppel, D., & Arnal, L. H. (2018). Proactive Sensing of Periodic and Aperiodic Auditory Patterns. *Trends in Cognitive Sciences*, *22*(10), 870–882. https://doi.org/10.1016/j.tics.2018.08.003

Rimmele, J. M., Poeppel, D., & Ghitza, O. (2021). Acoustically Driven Cortical δ Oscillations Underpin Prosodic Chunking. *Eneuro*, *8*(4), ENEURO.0562-20.2021. https://doi.org/10.1523/ENEURO.0562-20.2021

Rohenkohl, G., & Nobre, A. C. (2011). Alpha Oscillations Related to Anticipatory Attention Follow Temporal Expectations. *Journal of Neuroscience*, *31*(40), 14076–14084. https://doi.org/10.1523/JNEUROSCI.3387-11.2011

Romanovska, L., Janssen, R., & Bonte, M. (2019). Reading-Induced Shifts in Speech Perception in Dyslexic and Typically Reading Children. *Frontiers in Psychology*, *10*, 221. https://doi.org/10.3389/fpsyg.2019.00221

Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2012). Rhythm's gonna get you: Regular meter facilitates semantic sentence processing. *Neuropsychologia*, *50*(2), 232–244.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science (New York, N.Y.)*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Saksida, A., Iannuzzi, S., Bogliotti, C., Chaix, Y., Démonet, J.-F., Bricout, L., Billard, C., Nguyen-Morel, M.-A., Le Heuzey, M.-F., Soares-Boucaud, I., George, F., Ziegler, J. C., & Ramus, F. (2016). Phonological skills, visual attention span, and visual stress in developmental dyslexia. *Developmental Psychology*, *52*(10), 1503–1516. https://doi.org/10.1037/dev0000184

SanMiguel, I., Todd, J., & Schröger, E. (2013). Sensory suppression effects to self-initiated sounds reflect the attenuation of the unspecific N1 component of the auditory ERP. *Psychophysiology*, *50*(4), 334–343. https://doi.org/10.1111/psyp.12024

Sasisekaran, J., Smith, A., Sadagopan, N., & Weber-Fox, C. (2010). Nonword repetition in children and adults: Effects on movement coordination. *Developmental Science*, *13*(3), 521–532. https://doi.org/10.1111/j.1467-7687.2009.00911.x

Sato, M., & Shiller, D. M. (2018). Auditory prediction during speaking and listening. *Brain and Language*, *187*, 92–103. https://doi.org/10.1016/j.bandl.2018.01.008

Schaadt, G., & Männel, C. (2019). Phonemes, words, and phrases: Tracking phonological processing in pre-schoolers developing dyslexia. *Clinical Neurophysiology*, *130*(8), 1329–1341. https://doi.org/10.1016/j.clinph.2019.05.018

Scheerer, N. E., & Jones, J. A. (2014). The predictability of frequency-altered auditory feedback changes the weighting of feedback and feedforward input for speech motor control. *European Journal of Neuroscience*, *40*(12), 3793–3806. https://doi.org/10.1111/ejn.12734

Schiller, N. O. (2006). Lexical stress encoding in single word production estimated by event-related brain potentials. *Brain Research*, *1112*(1), 201–212. https://doi.org/10.1016/j.brainres.2006.07.027

Schiller, N. O., Jansma, B. M., Peters, J., & Levelt, W. J. M. (2006). Monitoring metrical stress in polysyllabic words. *Language and Cognitive Processes*, *21*(1–3), 112–140. https://doi.org/10.1080/01690960400001861

Schiller, N. O., Meyer, A. S., Baayen, R. H., & Levelt, W. J. M. (1996). A comparison of lexeme and speech syllables in Dutch. *Journal of Quantitative Linguistics*, *3*(1), 8–28. https://doi.org/10.1080/09296179608590060

Schmalz, X., Altoè, G., & Mulatti, C. (2017). Statistical learning and dyslexia: A systematic review. *Annals of Dyslexia*, *67*(2), 147–162. https://doi.org/10.1007/s11881-016-0136-0

Schmidt-Kassow, M., & Kotz, S. A. (2009). Event-related brain potentials suggest a late interaction of meter and syntax in the P600. *Journal of Cognitive Neuroscience*, *21*(9), 1693–1708.

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18. https://doi.org/10.1016/j.tins.2008.09.012

Schwartze, M., Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2011). Temporal regularity effects on pre-attentive and attentive processing of deviance. *Biological Psychology*, *87*(1), 146–151. https://doi.org/10.1016/j.biopsycho.2011.02.021

Serniclaes, W., Van Heghe, S., Mousty, P., Carré, R., & Sprenger-Charolles, L. (2004). Allophonic mode of speech perception in dyslexia. *Journal of Experimental Child Psychology*, *87*(4), 336–361.

Sheehan, K. A., McArthur, G. M., & Bishop, D. V. M. (2005). Is discrimination training necessary to cause changes in the P2 auditory event-related brain potential to speech sounds? *Cognitive Brain Research*, *25*(2), 547–553. https://doi.org/10.1016/j.cogbrainres.2005.08.007

Sitek, K. R., Mathalon, D. H., Roach, B. J., Houde, J. F., Niziolek, C. A., & Ford, J. M. (2013). Auditory Cortex Processes Variation in Our Own Speech. *PLoS ONE*, *8*(12), e82925. https://doi.org/10.1371/journal.pone.0082925

Snowling, M., & Hulme, C. (1994). The development of phonological skills. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *346*(1315), 21–27. https://doi.org/10.1098/rstb.1994.0124

Snowling, M. J. (2000). *Dyslexia*. Blackwell publishing.

Sowman, P. F., Kuusik, A., & Johnson, B. W. (2012). Self-initiation and temporal cueing of monaural tones reduce the auditory N1 and P2. *Experimental Brain Research*, *222*(1–2), 149–157. https://doi.org/10.1007/s00221-012-3204-7

Spratling, M. W. (2017). A review of predictive coding algorithms. *Brain and Cognition*, *112*, 92–97. https://doi.org/10.1016/j.bandc.2015.11.003

Sprugevica, I., & HOien, T. (2003). Early phonological skills as a predictor of reading acquisition: A follow-up study from kindergarten to the middle of grade 2. *Scandinavian Journal of Psychology*, *44*(2), 119–124. https://doi.org/10.1111/1467-9450.00329

Steinberg, J., Truckenbrodt, H., & Jacobsen, T. (2011). Phonotactic constraint violations in German grammar are detected automatically in auditory speech processing: A human event-related potentials study. *Psychophysiology*, *48*(9), 1208–1216. https://doi.org/10.1111/j.1469-8986.2011.01200.x

Stockall, L., Stringfellow, A., & Marantz, A. (2004). The precise time course of lexical activation: MEG measurements of the effects of frequency, probability, and density in lexical decision. *Brain and Language*, *90*(1–3), 88–94.

Storkel, H. L. (2009). Developmental differences in the effects of phonological, lexical and semantic variables on word learning by infants. *Journal of Child Language*, *36*(2), 291–321. https://doi.org/10.1017/S030500090800891X

Storkel, H. L., Armbrüster, J., & Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language, and Hearing Research: JSLHR*, *49*(6), 1175–1192. https://doi.org/10.1044/1092-4388(2006/085)

Storkel, H. L., & Lee, S.-Y. (2011). The Independent Effects of Phonotactic Probability and Neighborhood Density on Lexical Acquisition by Preschool Children. *Language and Cognitive Processes*, *26*(2), 191–211. https://doi.org/10.1080/01690961003787609

Summerfield, C., & de Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. *Nature Reviews. Neuroscience*, *15*(11), 745–756. https://doi.org/10.1038/nrn3838

Tamura, K., Karube, C., Mizuba, T., Matsufuji, M., Takashima, S., & Iramina, K. (2015). Phase-locked theta activity evoked in patients with severe motor and intellectual disabilities upon hearing own names. *Brain & Development*, *37*(8), 764–772. https://doi.org/10.1016/j.braindev.2014.11.009

ten Oever, S., Schroeder, C. E., Poeppel, D., van Atteveldt, N., & Zion-Golumbic, E. (2014). Rhythmicity and cross-modal temporal cues facilitate detection. *Neuropsychologia*, *63*, 43–50. https://doi.org/10.1016/j.neuropsychologia.2014.08.008

ter Bekke, M., Drijvers, L., & Holler, J. (2020). *The predictive potential of hand gestures during conversation: An investigation of the timing of gestures in relation to speech* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/b5zq7

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology*, *39*(4), 706–716. https://doi.org/10.1037/0012-1649.39.4.706

Thomas, E. R., Yon, D., de Lange, F. P., & Press, C. (2020). *Action enhances predicted touch* [Preprint]. Neuroscience. https://doi.org/10.1101/2020.03.26.007559

Thomson, J. M., Fryer, B., Maltby, J., & Goswami, U. (2006). Auditory and motor rhythm awareness in adults with dyslexia. *Journal of Research in Reading*, *29*(3), 334–348. https://doi.org/10.1111/j.1467-9817.2006.00312.x

Thomson, J. M., & Goswami, U. (2008). Rhythmic processing in children with developmental dyslexia: Auditory and motor rhythms link to reading and spelling. *Journal of Physiology-Paris*, *102*(1–3), 120–129.

Thorn, A. S. C., & Frankish, C. R. (2005). Long-term knowledge effects on serial recall of nonwords are not exclusively lexical. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*(4), 729–735. https://doi.org/10.1037/0278-7393.31.4.729

Tong, X., McBride, C., Zhang, J., Chung, K. K. H., Lee, C.-Y., Shuai, L., & Tong, X. (2014). Neural correlates of acoustic cues of English lexical stress in Cantonese-speaking children. *Brain and Language*, *138*, 61–70. https://doi.org/10.1016/j.bandl.2014.09.004

Tremblay, P., Deschamps, I., Baroni, M., & Hasson, U. (2016). Neural sensitivity to syllable frequency and mutual information in speech perception and production. *NeuroImage*, *136*, 106–121. https://doi.org/10.1016/j.neuroimage.2016.05.018

Van Atteveldt, N. M., Formisano, E., Blomert, L., & Goebel, R. (2007). The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. *Cerebral Cortex*, *17*(4), 962–974.

Van den Bos, K. P., & Lutje Spelberg, H. C. (2010). *Continu Benoemen en Woorden Lezen. Een test voor het diagnosticeren van taal- en leesstoornissen. Verantwoording*. Boom Test Uitgevers.

van Heuven, V. J., & de Jonge, M. (2011). Spectral and Temporal Reduction as Stress Cues in Dutch. *Phonetica*, *68*(3), 120–132. https://doi.org/10.1159/000329900

van Pelt, S., Heil, L., Kwisthout, J., Ondobaka, S., van Rooij, I., & Bekkering, H. (2016). Beta- and gamma-band activity reflect predictive coding in the processing of causal events. *Social Cognitive and Affective Neuroscience*, *11*(6), 973–980. https://doi.org/10.1093/scan/nsw017

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, *45*(1), 2–40. https://doi.org/10.1046/j.0021-9630.2003.00305.x

Verhagen, J., de Bree, E., Mulder, H., & Leseman, P. (2017). Effects of Vocabulary and Phonotactic Probability on 2-Year-Olds' Nonword Repetition. *Journal of Psycholinguistic Research*, *46*(3), 507–524. https://doi.org/10.1007/s10936-016-9448-9

Vidal, Y., Brusini, P., Bonfieni, M., Mehler, J., & Bekinschtein, T. A. (2019). Neural Signal to Violations of Abstract Rules Using Speech-Like Stimuli. *Eneuro*, *6*(5), ENEURO.0128-19.2019. https://doi.org/10.1523/ENEURO.0128-19.2019

Vitevitch, M. S., & Luce, P. A. (1998). When Words Compete: Levels of Processing in Perception of Spoken Words. *Psychological Science*, *9*(4), 325–329. https://doi.org/10.1111/1467-9280.00064

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic Phonotactics and Neighborhood Activation in Spoken Word Recognition. *Journal of Memory and Language*, *40*(3), 374–408. https://doi.org/10.1006/jmla.1998.2618

Vitevitch, M. S., & Luce, P. A. (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language*, *52*(2), 193–204. https://doi.org/10.1016/j.jml.2004.10.003

Vitevitch, M. S., Luce, P. A., Charles-Luce, J., & Kemmerer, D. (1997). Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech*, *40*(1), 47–62.

Wagner, R. K., Zirps, F. A., Edwards, A. A., Wood, S. G., Joyner, R. E., Becker, B. J., Liu, G., & Beal, B. (2020). The Prevalence of Dyslexia: A New Approach to Its Estimation. *Journal of Learning Disabilities*, *53*(5), 354–365. https://doi.org/10.1177/0022219420920377

Wang, H.-L. S., Huss, M., Hämäläinen, J. A., & Goswami, U. (2012). Basic auditory processing and developmental dyslexia in Chinese. *Reading and Writing*, *25*(2), 509–536. https://doi.org/10.1007/s11145-010-9284-5

Weber, C., Hahne, A., Friedrich, M., & Friederici, A. D. (2004). Discrimination of word stress in early infant perception: Electrophysiological evidence. *Brain Research. Cognitive Brain Research*, *18*(2), 149–161. https://doi.org/10.1016/j.cogbrainres.2003.10.001

Winkler, I., Denham, S. L., & Nelken, I. (2009). Modeling the auditory scene: Predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences*, *13*(12), 532–540. https://doi.org/10.1016/j.tics.2009.09.003

Wolpert, D. M., & Miall, R. C. (1996). Forward Models for Physiological Motor Control. *Neural Networks: The Official Journal of the International Neural Network Society*, *9*(8), 1265–1279. https://doi.org/10.1016/s0893-6080(96)00035-4

Wunderlich, J. L., & Cone-Wesson, B. K. (2001). Effects of stimulus frequency and complexity on the mismatch negativity and other components of the cortical auditory-evoked potential. *The Journal of the Acoustical Society of America*, *109*(4), 1526–1537. https://doi.org/10.1121/1.1349184

Ylinen, S., Shestakova, A., Huotilainen, M., Alku, P., & Näätänen, R. (2006). Mismatch negativity (MMN) elicited by changes in phoneme length: A cross-linguistic study. *Brain Research*, *1072*(1), 175–185. https://doi.org/10.1016/j.brainres.2005.12.004

Ylinen, S., Strelnikov, K., Huotilainen, M., & Näätänen, R. (2009). Effects of prosodic familiarity on the automatic processing of words in the human brain. *International Journal of Psychophysiology*, *73*(3), 362–368.

Yon, D., & Press, C. (2017). Predicted action consequences are perceptually facilitated before cancellation. *Journal of Experimental Psychology. Human Perception and Performance*, *43*(6), 1073–1083. https://doi.org/10.1037/xhp0000385

Zhang, M., Riecke, L., & Bonte, M. (2021). Neurophysiological tracking of speech-structure learning in typical and dyslexic readers. *Neuropsychologia*, *158*, 107889. https://doi.org/10.1016/j.neuropsychologia.2021.107889

Zora, H., Schwarz, I.-C., & Heldner, M. (2015). Neural correlates of lexical stress: Mismatch negativity reflects fundamental frequency and intensity. *Neuroreport*, *26*(13), 791–796. https://doi.org/10.1097/WNR.0000000000000426

Appendix

Impact Paragraph

The goal of the current research was to investigate how the brain uses familiar patterns of speech sounds (here: combinations of consonants) and speech rhythm (here: stress patterns) to process speech more efficiently. Every language has its own characteristic patterns in speech sounds and speech rhythm. For example, in Dutch the consonant combination -ts- is fairly common, while the combination -tk- is less common. The probability of different combinations of speech sounds is called phonotactic probability, so we could say that the pseudoword *notsal* has a high phonotactic probability compared to the pseudoword *notkal.* Similarly, first syllable stress is more common than second syllable stress in Dutch. Some theories suggest that implicit knowledge of these patterns can be used to anticipate or predict upcoming speech input. This can be particularly useful in situations where we cannot hear the speaker properly, such as when we are in an environment with background noise like a noisy bar, or when the sound on our Zoom call is disrupted because of poor internet connection. We know from previous studies that being able to learn these patterns is important for developing typical language and reading skills. Therefore, we also studied whether adults with dyslexia are able to use these familiar patterns in the same way as typical readers.

The current research used electroencephalography (EEG), which is a non-invasive method of measuring electrical brain activity through electrodes that are placed on the scalp. Our main findings highlight that familiar patterns of speech sounds (as in the pseudoword *notsal*) allow native Dutch speakers to process speech more efficiently, but that this effect is reduced in dyslexic adults. In contrast, differences in stress patterns (first vs. second syllable stress) did not appear to influence how Dutch speakers process pseudowords presented in isolation. This dissertation has also highlighted that changes in sound combinations and speech rhythm are processed with different brain mechanisms, which we can observe through different changes in neural oscillations. Finally, we show preliminary evidence suggesting that regularities influence speech processing differently depending on whether the speech is self-generated or externally presented.

By studying dyslexic readers, this dissertation extends our understanding of reading difficulties to inform research leading to future interventions. Developmental dyslexia is characterized as persistent reading difficulties despite adequate schooling. Reports of the prevalence of dyslexia range from 3% to values as high as 20%,

however, these estimates vary due to differences in diagnostic criteria across countries and institutions, as well as different characteristics of languages such as orthographic transparency. Developmental dyslexia can have a substantial impact on the individual later in life, as it is associated with poor outcomes in a range of domains including academic and professional success as well as emotional wellbeing. Thus, early diagnosis and intervention is crucial for improving later life outcomes. Recent theories propose atypical processing of rhythm may be an underlying risk factor for various developmental language difficulties, such as dyslexia and specific language impairment. While we were unable to find evidence for atypical processing of speech rhythm in Dutch adult dyslexic readers, previous studies have shown differences in neural sensitivity to variations in syllable stress in pre-reading children who later develop reading difficulties. Future approaches could apply the paradigm used in Chapters 2 and 3 to pre-reading children, to study differences in the developmental trajectory of a sensitivity to syllable stress regularities and phonotactic probability and their relationship with phonological, rhythmic and reading skills at different stages of development. Such investigations could eventually give insights into potential early markers for dyslexia, to allow for early detection and intervention.

The research in this thesis has been presented at several international scientific conferences to contribute to scientific exchange and progress in the field. All findings reported in this thesis either have been or will be made publicly available via publication in open access journals to support accessibility of the generated knowledge. In addition to communicating with the scientific community, aspects of the current research have also been shared with the general public in the Netherlands, Germany and the United States, including introductions on how to study the brain for children and adolescents in primary and secondary schools as well as talks for bachelor students in linguistics to introduce them to the interdisciplinary field of studying language in the brain, and online events highlighting the work of women in science at Maastricht University. To promote accessibility of science communication to non-English speaking communities, the research has also been shared with a German-speaking audience.

Appendix

Curriculum Vitae

Alexandra (Alex) Katherine Emmendorfer was born on April 11th, 1991 in New Haven, Connecticut, USA. In January 1999, her family relocated to Erding, Germany for 3 years*. She completed her secondary education at the Anne Frank Gymnasium Erding, finishing her Abitur in May 2011. Alex returned to the United States to enroll in a bachelor program at Northeastern University, where she studied Behavioral Neuroscience with a minor in Linguistics, with the intention of later attending medical school to become a neurologist. Through Northeastern's Co-operative Education program, Alex was able to complete several research internships during her bachelor studies. During her first internship, where she studied visual processing in mice at KU Leuven in Belgium, she made three decisions: (1) to return to Europe after her bachelor studies, (2) to pursue a career in research instead of medicine but (3) not on visual processing in mice. After exploring her research interests in further internships and graduating from Northeastern University magna cum laude in 2015, she followed through with those decisions by enrolling in the Research Master in Cognitive Neuroscience at Maastricht University. For her master thesis, she investigated formal and temporal predictions in speech perception, which paved the way for her subsequent PhD project. Alex graduated cum laude from the Research Master in Cognitive Neuroscience in 2017, and subsequently began her PhD under the supervision of Milene Bonte, Bernadette Jansma and Sonja Kotz. She is now working under Judith Holler as a postdoctoral researcher at the Donders Centre for Cognition investigating the role of visual communication signals (e.g., gestural & facial signals) in language processing.

*As of writing this nearly 23 years later, her parents and brother are still in Erding, Germany, with no intentions of leaving.

Appendix

List of Publications

Peer-reviewed publications

Emmendorfer, A. K., Correia, J. M., Jansma, B. M., Kotz, S. A., & Bonte, M. (2020). ERP mismatch response to phonological and temporal regularities in speech. *Scientific reports*, *10*(1), 1-12

Santarnecchi, E., **Emmendorfer, A.,** Tadayon, S., Rossi, S., Rossi, A., & Pascual-Leone, A. (2017). Network connectivity correlates of variability in fluid intelligence performance. *Intelligence*, *65*, 35-47.

Santarnecchi, E., **Emmendorfer, A.,** & Pascual-Leone, A. (2017). Dissecting the parieto-frontal correlates of fluid intelligence: A comprehensive ALE meta-analysis study. *Intelligence*, *63*, 9-28.

Sprugnoli, G., Rossi, S., **Emmendorfer, A.,** Rossi, A., Liew, S. L., Tatti, E., ... & Santarnecchi, E. (2017). Neural correlates of Eureka moment. *Intelligence*, *62*, 99-118.

Preprint articles

Emmendorfer, A. K., Bonte, M., Jansma, B. M., & Kotz, S. A. (2021). Phonological and temporal regularities lead to differential ERP effects in self-and externally generated speech. *bioRxiv*.

Manuscripts in preparation

Emmendorfer, A. K., Kotz, S. A., Jansma, B. M., & Bonte, M. (In Preparation). ERP and time-frequency correlates of phonological and temporal deviants in dyslexic readers.

Conference contributions

Emmendorfer, A. K., Kotz, S. A., Jansma, B. M., & Bonte, M. (2021, October). *ERP and time-frequency correlates of phonological and temporal deviants in dyslexic readers.* Slide Slam presentation at Society for Neurobiology of Language, 2021. (Virtual).

Emmendorfer, A. K., Jansma, B. M., Kotz, S.A., & Bonte, M. (2020, October). *Atypical processing of phonotactic probability and syllable stress in dyslexic adults: an MMN study.* Poster presented at Society for Neurobiology of Language 2020. (Virtual).

Emmendorfer, A. K., Jansma, B. M., Kotz, S.A., & Bonte, M. (2019, December). *Atypical processing of phonotactic probability and syllable stress in dyslexic adults: an MMN study.* Poster presented at Rate and Rhythm in Speech Recognition, Nijmegen, The Netherlands.

Emmendorfer, A. K., Bonte, M. L, Jansma, B. M., & Kotz, S. A. (2019, August). *Motor-induced suppression of N1 and P2 is modulated by phonotactic probability and syllable stress.* Poster presented at Society for Neurobiology of Language 2019. Helsinki, Finland.

Emmendorfer, A. K., Correia, J., Schroen, J., Jansma, B. M., Kotz, S. A., & Bonte, M. L. (2018, August). *Formal and temporal predictions in speech perception.* Poster presented at Society for Neurobiology of Language 2018. Québec City, Canada.

Emmendorfer, A., Correia, J., Kotz, S., & Bonte, M. (2017, June). *Formal and temporal predictions in speech perception.* Poster presented at Dutch Neuroscience Meeting 2017, Lunteren, The Netherlands.

Santarnecchi, E., Fox, M. D., Almquist, J., Brem, A. K., Cohen Kadosh, R., Dillard, M., **Emmendorfer, A.**, (...) Pascual-Leone, A. (2015, May). *Predicting the outcome of computerized cognitive training using resting-state functional connectivity patterns.* Poster presented at ESCoN 2015, San Fransisco, CA.

Presentations and outreach

Emmendorfer, A. K. (March, 2021). *Studying the Brain with EEG.* Invited talk to high school students. At College den Hulster, Venlo, Netherlands. (Virtual)

Emmendorfer, A. K. (March, 2021). *(Ir)regularities in speech perception.* At: Women Researcher's Festival – International Women's Day, organized by FEM Maastricht and Pint of Science Maastricht (Virtual)

Emmendorfer, A. K. (November, 2020). *From Linguistics to Cognitive Neuroscience: Studying the Brain from a Linguist's Perspective.* Invited talk to undergraduate students of the Linguistics Program at Northeastern University. Boston, USA. (Virtual)

Emmendorfer, A. K. (November, 2019). *Das Gehirn blickt in die Zukunft: wie (Un)Regelmäßigkeiten in der Sprache zur Sprachverarbeitung beitragen.* Invited talk at Volkshochschule Erding. Erding, Germany.

Emmendorfer, A. K. (November, 2016). *Studying the brain's role in listening, speaking, and reading: how do predictions help us learn?* Skype session with 6th grade Science class at The Neighborhood House Charter School. Dorchester, USA

Appendix

Acknowledgements

186

The work carried out in this thesis would not have been possible without the help of many colleagues, family, and friends who supported me along the way. Milene, Sonja and Bernadette, or more simply: Team Milsonette. Thank you for believing in me and giving me the opportunity to do this PhD. I am particularly grateful for the support you gave during some of the most challenging moments of my PhD, in the midst of a global pandemic, when I was beginning to doubt everything. Our Monday morning meetings gave me the structure I needed to get through the week and work towards finishing the thesis step by step, and I do not take this support for granted.

Milene: your strength and leadership are truly inspiring. Thank you for always keeping a clear head during the rocky periods.

Sonja: I am immensely grateful for the atmosphere you have created in the BAND-Lab, fostering support and caring both in and outside the lab.

And finally, Bernad**ette**: who has already been my mentor since the research master. I am thankful that I could always count on your warm smile and encouragement during my time in Maastricht. While not included in the official list, I think a central proposition that came out of my PhD is that everyone needs a Bernie in their life. I am grateful to have you on my team.

João, you were the best master thesis supervisor I could have asked for. I am so grateful for the supervision and guidance you provided during my internship. You managed to hit the sweet spot of leaving room for me to learn how to solve problems on my own, while offering enough support so I wouldn't feel completely lost. Thank you for helping me develop the skills that were so crucial for my PhD!

Miriam, my better-CN-half, bereits seit dem Research Master. Wir haben so vieles zusammen gelernt: receptive fields & tonotopic maps, personal growth, and making tempeh taste good. Ich bin froh dass wir diese PhD Zeit zusammen erleben konnten. Du warst immer da um mir die selbstverständliche Lösung zu scheinbar unüberwindbaren Problemen zu zeigen. Über die Jahre haben wir zusammen gekocht, getanzt, gelacht, geweint, und Gesaffelstaffeln geguckt. Ich bin für jeden Moment dankbar, und stolz dich als Freundin zu haben.

The BAND-Lab ladies. I have always been happy to have you as colleagues and friends, but I am especially grateful for the support network we have now created for each

other since early 2020, including lunch walks, Zoom office sessions, and a group chat full of support, love, and affirmations. If there is anything positive to take away from this period, it is that it has brought us closer together.

First and foremost, Lisa and Xan: Thank you for letting me work (and sometimes nap) on your office couch so I didn't have to run back and forth between buildings on testing and meeting days. Thank you for always being available to rant and complain, but also to share happy moments like hitting "submit". I always look forward to our coffee Zooms where Lisa grills us not only on *what* we are cooking for dinner, but also *how* we are cooking it. And finally, thank you for the donuts!

Katerina. In one of our early heart-to-hearts, you gave me a piece of advice that I have deeply cherished: I am the expert on my project. Other people may know more about a lot of other things, but no one knows my project better than me. This gave me the strength to get through many talks and poster presentations, and around the time you read this, hopefully also my PhD defense. Thank you for always being a source of inspiration, both in the lab and in life.

Pia & Hanna. It has been so inspiring to see how you tackle all the challenges that come your way, while keeping a positive attitude at the end of the day. I am so excited to see what is in store for you!

And of course, many thanks to the other past and present members of the BAND-Lab and MBIC Language Group who have shaped these years: Michael, Francesco, Laura, Rachel, Gojko, Joe, Linda, Manli, Giada and Antonio. Thank you for your support, thoughtful comments, and critical questions during presentations, and most importantly, thank you for all the delicious dinners together!

Lars, thank you for taking the time to help me solve my data problems. Your support was so valuable in helping me regain a bit of confidence.

Thank you to the students and interns who have helped me with data collection: Alexa, Anna, CJ, Dana, Gughi, Helena, Venja. Your help in the lab made data collection run smoothly. Thank you for helping me prepare, fit, and clean over 100 EEG caps, and for keeping me company in the lab on long testing days so it is not so lonely!

Heather. You taught me how to write a paper and instilled a curiosity in me for all things related to language, which I simply haven't been able to shake. Thank you for your support and encouragement during my time at Northeastern, for taking me on in the research seminar even though I was "only" minoring in linguistics, and for being a truly inspiring instructor and role-model.

What do you get when you cross a software engineer and an English teacher? Apparently one possible outcome is a cognitive neuroscientist with an interest in language. On that note, I would like to thank my family: While I always stubbornly fought to be independent, starting at a young age, I would not be where I am today if it weren't for my parents' support. I am so grateful that you took the brave decision in 1999 to uproot your home and move across the Atlantic Ocean with 2 kids and 2 cats, to a foreign country where neither of you spoke the language. Having done this myself with no kids, only 1 cat, and Google Translate, I recognize that this decision took a lot of courage. Without a doubt, it had a great impact on our lives, and I am glad that I had my brother Nick to go through this experience with. I can't imagine what we would be doing now if we had stayed in the US.

Marla, James, Justine, Patrick and Ellie. Thank you for being my home away from home while I lived in Boston. I hope I can one day return the favor for Justine and Patrick. I am glad I had the opportunity to watch them grow up during those years. I am so proud of them!

Merci à ma famille française: Odile, Alain, Anne, Isabelle, Catherine, Mado et Clément. Pendant les dernières années il n'était pas possible de vous rendre visite, hélas. Néanmoins, je suis contente de savoir que j'ai toujours une place à la campagne française pour pouvoir profiter de la vie à l'air frais, du bon vin, de la bonne nourriture, mais le plus important, de pouvoir en profiter avec vous.

The remaining Emmendorfer and Redinger clans scattered across the US: thanks for your love and support from far away. Mémé, Pépé, and Grandma, I miss you and I wish you were here today.

Domo. Thank you for your patience and support, even if you don't always understand all the dynamics of my work and sometimes doubt whether I have a "real job". I can always count on you to remind me that there is life outside of my academic bubble,

and to help me switch off from work at the end of the day by distracting me with strange hypothetical D&D scenarios and new character ideas. Thank you for introducing me to climbing and the Avatar series. Thank you for taming my cat and for helping me (re)learn how to drive. And finally, thank you for always supporting me to make decisions for myself, and for encouraging me to face my fears. In 4 years of PhD, we've successfully veganized several traditional Hungarian dishes (nagyon finom!), shared nearly 200 homemade pizzas, and almost as many silly nicknames. I hope to continue these rituals with you for many years to come.

Having some creative outlets and people to share them with was crucial to maintaining a balance during this PhD. For me these outlets were cooking, climbing, and D&D, and they were all made more enjoyable because of the people I shared them with. Dale, Frea, Jana, Maike, Manoj, Sally, Sven, Thijs, and the many past and present Maastricht Goes Vegan team members and climbing buddies. Thank you for the delicious meals together, either Tuesday evenings at Maastricht Goes Vegan, or weekends at home. Thank you for the hours cheering each other on and solving difficult problems together at the boulder gym. Thank you for regular escapes from reality in the stories we are creating with Mara, Fenniana, Drous, Frank, Carmelata, Clorck, Nova, Jocko and many more (in particular, shout-out to Domo and Sally for the hard work you put into DMing).

Nadja, auch wenn wir uns in den letzten Jahren nicht oft sehen konnten, bin ich für deine Freundschaft und Unterstützung dankbar. Es freut mich so zu sehen wie gut es dir mit deiner kleinen Familie geht!

Neil Dave? Daveh? Somehow, despite our completely crazy schedules preventing us from ever spending time together while I lived in Boston, you were someone I could always count on when it mattered. Thanks for continuing to be that person, even with an ocean between us.

It is no secret that there is a high prevalence of depression and anxiety among PhD candidates (e.g., Satinsky et al., 2021), and so it may not come as a surprise that my mental health has suffered in the process of creating this thesis. A crucial part of making it to the end was the support of my friends, colleagues, and family who I have already thanked above. I'd also like to thank Anna, for creating a platform to talk

openly about academic mental health, as well as the other co-founders and members of Flourish Maastricht, who are working towards improving the mental health support for students and staff at Maastricht University. Thank you to Tessa, for the workshops offered via the staff development center, and thank you to the two other PhD candidates who participated in her workshop "Overcome your fear of failure" with me. I am glad that we were able to stay in touch after the end of the workshop to continue to support each other! Thank you to the mental health care professionals who have supported me over these last years, in particular Daisy, Shannen and Iris. And finally, I would like to thank *myself* for having the courage to take on this challenge, and the strength to make it to the finish line. In sharing this here, I hope that I can do my little part in reducing the stigma of talking about mental health and make it a little bit easier for anyone reading this to seek out the support they need.

Satinsky, E. N., Kimura, T., Kiang, M. V., Abebe, R., Cunningham, S., Lee, H., ... & Tsai, A. C. (2021). Systematic review and meta-analysis of depression, anxiety, and suicidal ideation among Ph. D. students. *Scientific Reports*, *11*(1), 1-12.