Donders Graduate School for Cognitive Neuroscience
Master of Science Programme
MSc Thesis

Spectrotemporal Modulation Sensitivity in Developmental Dyslexia

Yingdi Xie[1]

Supervisors: Kiki Van der Heijden[2,3], Elia Formisano[2,3], Milene Bonte[2,3]

[1]Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, The Netherlands
[2]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands
[3]Maastricht Brain Imaging Centre (MBIC), Maastricht, The Netherlands

Abstract

Previous research has suggested various general auditory processing deficits which may underlie the reduced phonological awareness in developmental dyslexia. However, despite the importance of spectrotemporal modulations for speech processing, there is no study to date which systematically examined auditory processing of the modulation components characteristic of speech in dyslexia. Thus, the present study aims to address if dyslexic and normal readers differ in perceptual sensitivity to these spectrotemporal modulations. We predict a reduced sensitivity in dyslexic readers. We used adaptive transformed up-down procedure (Chi et al., 1999; Levitt, 1971) to estimate detection thresholds of dyslexic and normal readers for different combinations of spectrotemporal modulations in dynamic ripples and AM broadband noises. Contrary to our prediction, multilevel modeling revealed that there was no significant group difference, indicating comparable modulation sensitivity between dyslexic and normal readers. It opposes all present hypothesized auditory deficits. Moreover, we found a significant interaction between the effects of temporal modulations and those of spectral modulations. It implies a dependency of these two processing mechanisms. Future research is needed to further inspect the auditory processing of speech as well as other natural sounds in dyslexia.

Keywords: spectrotemporal modulations, developmental dyslexia, auditory perception, psychophysics

Corresponding author: Yingdi Xie, e-mail: xieyd.qy@gmail.com

1. Introduction

Developmental dyslexia is a persistent deficit in reading acquisition despite normal general intelligence and adequate education opportunity (Lyon, 1995), which affects approximately 7% of children (Goswami, 2011). In addition to poor reading skills, dyslexic readers exhibit reduced phonological awareness, which might be due to an imprecise representation of phonemes in the brain (Elbro, Borstrøm, & Petersen, 1998). Specifically, dyslexic readers have difficulties with decoding and manipulating phonemes (Stanovich, 1988), as is revealed by phonological awareness tasks such as *spoonerisms* which requires subjects to exchange the beginning sounds of two words (e.g., "Walt Disney" should become "Dalt Wisney"; Snowling, Nation, Moxham, Gallagher, & Frith, 1997). Moreover, phonological awareness reliably predicts subsequent reading ability (Lundberg, Olofsson, & Wall, 1980). Despite this clear characterization of the core deficit in developmental dyslexia, the underlying developmental mechanisms are still largely unknown. Several hypotheses have been proposed, such as the magnocellular theory (Stein & Walsh, 1997) and the cerebellar deficit hypothesis (Nicolson, Fawcett, & Dean, 2001). Interestingly, one line of research suggests that dysfunctional auditory processing may lead to distorted speech perception and thereby underlie the reduced phonological awareness in dyslexic readers. In the subsequent sections, we will first discuss the acoustic characteristics of speech, in particular spectrotemporal modulations and their role in speech processing, and then turn to an overview of the hypothesized auditory processing deficits in dyslexia and their effects on speech perception.

1.1 Speech and spectrotemporal modulations

Speech features fluctuation of acoustic energy in the temporal and spectral dimensions, that is, spectrotemporal modulations. The energy is not evenly distributed but mostly present in slow modulation rates. For instance, an acoustic analysis of Dutch speech using a bank of modulation filters showed that the power concentrates at < ~16 Hz for temporal modulations (Figure 1A,B) and at < ~3 cycles/octave for spectral modulations (Figure 1A,C), with downward sweeps (i.e., positive temporal modulations) being slightly stronger than upward ones (Figure 1A,B), in line with the findings on English speech by Singh and Theunissen (2003). It is also noticeable that speech energy is mainly restricted to a low-frequency range (< ~1 kHz) (Figure 1B,C), consistent with the analysis of Goswami (2015).

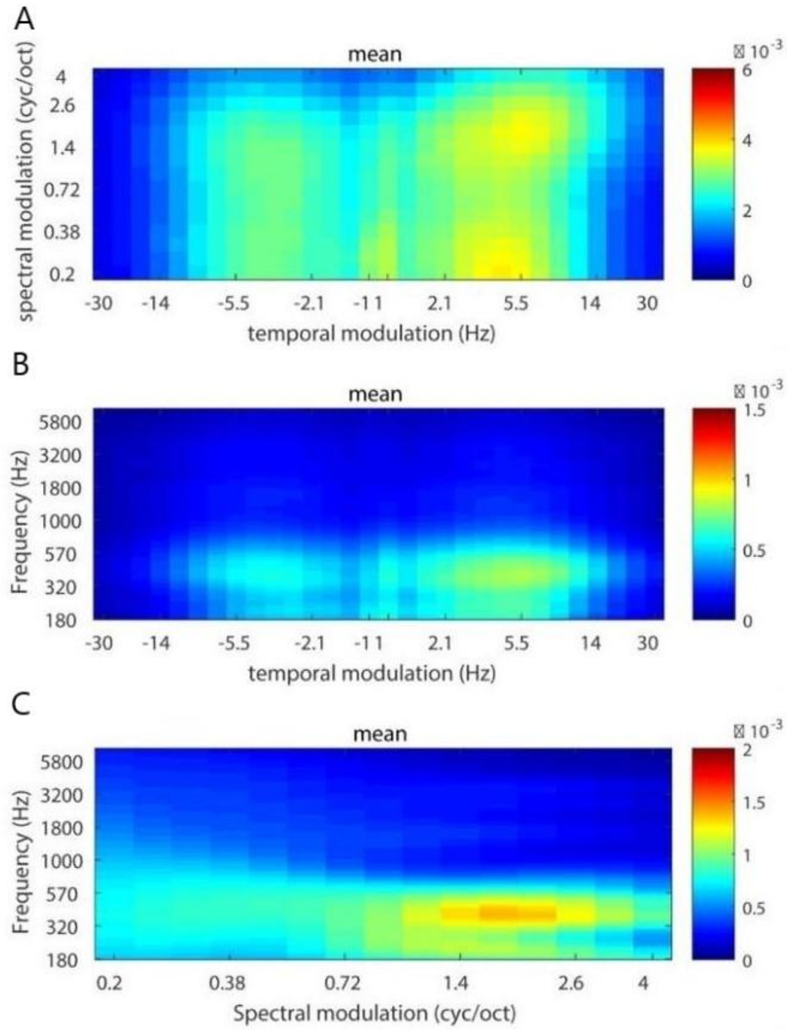

Figure 1. Spectrotemporal modulation spectra of a set of Dutch speech sentences. Time-averaged power is plotted in the color scale, and the other dimensions (spectrotemporal modulations and frequency) on the axes. Specifically, each panel plots two dimensions (spectrotemporal in A, frequency and temporal in B, and frequency and spectral in C), with the power summed across the third dimension. By courtesy of Van der Heijden.

Importantly, it has been shown that spectrotemporal modulations are vital for speech intelligibility and that smearing them will prevent speech or phonemic processing. For instance, Drullman, Festen and Plomp (1994) found that the sentence intelligibility is severely reduced by removing temporal modulations from 4 to 16 Hz. Liu and Eddins (2008) showed that filtering out spectral modulations below 2 cycles/octave degrades accuracy in vowel identification. In addition, a systematic investigation using notch filters revealed that erasing temporal energy from 1 to 7 Hz and spectral energy below 1 cycle/kHz significantly impairs speech intelligibility (Elliott & Theunissen, 2009).

Interestingly, the human auditory system seems to be particularly sensitive to the spectrotemporal modulations which are crucial for speech. Classic psychoacoustic experiments measured detection thresholds for a range of spectrotemporal modulations. These detection thresholds can be translated into a spectral and temporal modulation transfer function (sMTF and tMTF, respectively). The shape of the tMTF resembles a low-pass filter (Chi, Gao, Guyton, Ru, & Shamma, 1999; Viemeister, 1979), with perceptual sensitivity peaking around 1-8 Hz and decreasing at faster temporal rates. The sMTF resembles a band-pass filter (Eddins & Bero, 2007) or low-pass (Chi et al., 1999; Green, 1986), indicating a reduced sensitivity as the spectral modulation rate increases. Although there is some inconsistency in the precise shape of the MTFs, it is commonly demonstrated that humans have a higher sensitivity for slow spectrotemporal modulations which are critical for speech comprehension.

Studies using neurophysiological and neuroimaging techniques showed that neural processing also appears to reflect this increased sensitivity for slow spectrotemporal modulations that are predominant in speech. For example, several functional magnetic resonance imaging (fMRI) studies showed that human non-primary auditory cortex displays the strongest response to slow temporal modulations (~4-8 Hz) in AM broadband noises (Giraud et al., 2000), AM narrowband noises (Overath, Zhang, Sanes, & Poeppel, 2012) or natural sounds (Santoro et al., 2014). Moreover, these temporal modulations are more preferably encoded in the brain than faster rates. For instance, by employing model-based stimulus reconstruction approach on electrocorticography (ECoG) signals, Pasley et al. (2012) showed that slow and intermediate rates (< ~8 Hz) are more accurately reconstructed than faster ones and that they can be faithfully decoded with a linear model of spectrogram, whereas fast fluctuations require a nonlinear model based on modulation energy. Similarly, Santoro et al. (2017) found that low-frequency temporal modulations (~2-4 Hz) can be more accurately decoded from measured fMRI signals than higher rates. Critically, this result was not affected by the exclusion of speech stimuli from the decoding analysis, suggesting that the measured responses reflect low-level processing of acoustic characteristics, rather than higher-order processing of speech. In addition, slow temporal modulations are pertinent for the neural oscillation network model of speech perception (Giraud & Poeppel, 2012; Kösem & van Wassenhove, 2017; Luo & Poeppel, 2007). This model argued that the phase-locking (or entrainment) of neural oscillatory activities to the delta/theta-band (~1-

8 Hz) temporal modulations in speech guides the segregation and analysis of linguistic information, thus lying in the pivotal process of cortical speech processing.

In short, slow spectrotemporal modulations are strongly present in speech, are essential for speech comprehension, and neural auditory processing is optimized for processing these modulations, that is, for processing the features characteristic of speech.

1.2 Dyslexia and general auditory perception

Previous research has examined diverse aspects of general auditory perception in dyslexia. This section will review various proposed auditory processing deficits and how they affect the speech perception of dyslexic readers. We will first address the hypotheses on the processing of slow spectrotemporal modulations and then other views on the perception of fast components in speech.

As discussed before, slow spectrotemporal modulations are predominant acoustic characteristics of speech and play a crucial role in speech comprehension. It is thus hypothesized that abnormal processing of these modulations may underlie dyslexia. One example is the amplitude modulation deficit hypothesis (Goswami et al., 2002), which argued that the slow amplitude envelope (i.e. slow temporal modulations) is processed less accurately in dyslexic readers. Specifically, the deficit resides in detecting the onset or rise time of the speech envelope (Goswami et al., 2002). That is, the acoustic contour at a modulation rate of around 2-4 Hz provides rhythmic cues that facilitate syllable segmentation. Consequently, a general difficulty in tracking the amplitude envelope will lead to defective parsing and thereby affect the phonological encoding of words. (Goswami, 2011, 2018) Supporting evidence arises from research revealing that by contrast with normal readers, dyslexic readers are significantly poorer at discriminating the syllabic stress pattern (Leong, Hämäläinen, Soltész, & Goswami, 2011), detecting the gap in otherwise continuous tones (Trehub & Henderson, 1996), and accurately tapping along to a rhythmic beat (Leong & Goswami, 2014). In addition, detection thresholds for temporal modulations are higher in children with developmental dyslexia than in normally reading children at a modulation rate of 4 Hz (Lorenzi, Dumont, & Füllgrabe, 2000).

Neurophysiological research findings also support this theory. For instance, event-related potentials (ERP) in response to AM noise are significantly weaker in dyslexic readers than normal readers (Menell, McAnally, & Stein, 1999). Moreover, the amplitude of the auditory ERP

P1 component is reduced in dyslexic readers when the rise time of envelope is extended, which suggests an insensitivity to the envelope onset (Stefanics et al., 2011). Moreover, in light of the neural oscillation network model of speech perception (Giraud & Poeppel, 2012), the phase-locking (or neural entrainment) to speech envelope plays a central role in speech analysis. Indeed, it was found that dyslexic readers entrain less accurately or at a longer phase lag than normal readers in the delta-band (~2 Hz) oscillations (Abrams, Nicol, Zecker, & Kraus, 2009; Stefanics et al., 2011).

In addition to the processing of slow temporal modulations, some studies showed that dyslexic readers may also have difficulties with perceiving slow spectral modulations characteristic of speech. For instance, psychophysical experiments using frequency-modulated (FM) tones found that detection thresholds for spectral modulations at 2 Hz are higher in dyslexic children (Talcott et al., 2000) and adults (Ramus et al., 2003) than in normal counterparts. Furthermore, Boets et al. (2011) showed that the detection thresholds in kindergarten children with a family history of dyslexia predict their reading and spelling ability in the first grade of school.

In opposition to the hypothesized auditory processing deficit in slow spectrotemporal modulations, some theories suggest that the impairment lies in the processing of fast components, such as the rapid auditory processing deficit hypothesis (Tallal, 1980). Specifically, Tallal (1980) argued for a deficiency in processing brief, rapidly successive auditory stimuli. That is, speech contains transient changes in spectral profile, often lasting for merely tens of milliseconds, which distinguish phonological contrasts (e.g., the onsets of *ba* vs.. *da*). Hence, a failure to perceive fast frequency transitions will result in an unspecified or even incorrect phonemic representation. This in turn affects grapheme-to-phoneme conversion, a core stage of reading development. (Tallal, 1980)

Evidence for this theory derives from the finding that dyslexic readers are less able than normal readers to discriminate tones or to repeat temporal pitch patterns when the inter-stimulus interval (ISI) is small, while their performance is normal at longer ISIs (Tallal, 1980). Subsequent studies (Ahissar, Protopapas, Reid, & Merzenich, 2000; Farmer & Klein, 1995; Reed, 1989) replicated the results. In addition, the discrimination accuracy of dyslexic readers is lower than that of normal readers when the frequency of the tones varies (Ahissar et al., 2000). Moreover, studies indicated that dyslexic readers have difficulty discerning phonological contrasts separated by a short ISI, or in time-compressed speech (Watson, Stewart, Krause, & Rastatter, 1990), particularly those contrasts of brief stop consonants than long vowels (Reed, 1989), as would be predicted by the rapid auditory processing deficit hypothesis.

Besides the perception of spectral transients in speech, some studies showed that the deficit may exist in processing fast temporal modulations. For example, Menell et al. (1999) found that dyslexic readers have reduced sensitivity to the temporal modulations ranging from 10 to 320 Hz. Moreover, in the tMTF of Lorenzi et al. (2000), dyslexic readers display a higher detection threshold for a modulation rate at 1024 Hz than normal readers. However, how the defective perception of these fast temporal modulations affects speech perception is still unspecified.

In brief, although these studies tested different aspects of auditory perception in dyslexia, they all hint at a causal link to deviant processing of temporal and/or spectral components. Yet, there are several caveats. First, a number of experiments failed to replicate effects. For instance, in contrast to the hypothesized rapid temporal processing deficit, some studies showed that dyslexic readers are not harmed by short ISIs while performing frequency discrimination (Amitay, Ahissar, & Nelken, 2002; Goswami, Fosker, Huss, Mead, & Szűcs, 2011), and that stretching formant transitions in either frequency or time domain does not help dyslexic readers discern phonological contrasts (McAnally, Hansen, Cornelissen, & Stein, 1997). Regarding the amplitude modulation deficit hypothesis, some studies also showed comparable detection of envelope rise time in dyslexic and normal readers (Amitay et al., 2002; Georgiou, Protopapas, Papadopoulos, Skaloumbakas, & Parrila, 2010).

Second, the selection of acoustic parameters is often unclear and varies widely across studies. For example, Lorenzi et al. (2000) and Menell et al. (1999) both tested temporal modulation sensitivity but used a different range of modulation rates of which most are not prominent in speech.. To the best of our knowledge, there is no study that systematically examined auditory processing across the full spectrotemporal modulation space, and in particular of the modulation components characteristic of speech.

Finally, most studies used simple sound stimuli, such as pure tones (e.g., Ahissar et al., 2000), AM broadband noises (e.g., Lorenzi et al., 2000) or FM tones (e.g., Talcott et al., 2000). However, as discussed previously, speech and phonemes are characterized by multi-dimensional features including concurrent modulations of temporal and spectral envelopes. Hence, it may not be possible to generalize the findings of the studies employing relatively simple stimuli to the domain of speech processing.

1.3 The current study

The current project aims to investigate the proposed auditory processing deficits in a comprehensive study that systematically evaluates the sensitivity of dyslexic readers to different spectrotemporal modulations. Specifically, the present study addresses the following question: do dyslexic and normal readers differ in perceptual sensitivity to spectrotemporal modulations in complex sounds?

We use a classic psychophysical paradigm (Chi et al., 1999; Levitt, 1971) with dynamic ripples (i.e., broadband noises which are simultaneously modulated in the temporal and spectral dimensions) as well as AM broadband noises to estimate the detection thresholds for different combinations of spectral and temporal modulation rates in dyslexic readers as well as normal readers. This paradigm has been used by Chi et al. (1999) to estimate modulation detection thresholds and derive tMTF and sMTF in normal population and we thereby adapt from their study using similar parameters to test dyslexic readers.

We hypothesize that dyslexic readers have a higher threshold for the spectrotemporal modulations which are important for speech (e.g., <~8 Hz in temporal modulations), suggesting a reduced perceptual sensitivity to the spectrotemporal modulations most prominent for speech. Thus, we predict a main effect of group (dyslexic vs. normal readers), or a significant interaction between group and the effects of spectral/temporal modulations, indicating a (modulation-specific) difference in perceptual sensitivity. In addition, we expect a main effect of temporal and spectral modulation rates. That is, we expect different detection thresholds depending on the spectrotemporal modulation rate (e.g., Chi et al., 1999).

2. Methods

2.1 Participants

Ten officially-diagnosed developmental dyslexic readers (4 female, mean age = 20.6 years, ranging from 19.1 to 40.0 years) and sixteen normal readers (9 female, mean age = 23.2 years, ranging from 19.6 to 35.1 years) were recruited for our study. All were Dutch native speakers and the two groups were matched on education level and linguistic background. Participants gave permission by voluntarily signing a consent form before start and received monetary compensation in the end. One normal reader was excluded from data analysis in that he failed to detect modulations at the largest depth (100%) in 3 out of 15 conditions, showing an outlying performance. Thus, a total control group being analyzed consisted of 15 participants (9 female, mean age = 22.5 years, ranging from 19.6 to 27.8 years). Participants had normal hearing as assessed with standard audiometric testing of pure-tone hearing thresholds (0.25, 0.5, 1, 2, 4 and 6 kHz) using a standard audiometer (MAICO MA30). The Ethical Committee of the Faculty of Psychology and Neuroscience at Maastricht University granted approval for our study.

2.2 Materials

Sound stimuli were created using MATLAB (2014a; The MathWorks, Natick, 2014). We used similar acoustic parameters as described in Chi et al. (1999). That is, the standard reference was broadband noises consisting of 92 tones equally spaced along the logarithmic frequency axis, ranging from 140 to 7340 Hz and spanning 5.75 octaves. For the targets (dynamic ripples), we selected 5 different temporal rates (2, 4, 8, 16 and 32 Hz) and 2 different spectral rates (1 and 2.5 cycles/octave) which cover the spectrotemporal modulation space characteristic of speech. In addition, we chose a pure temporal modulation condition (0 cycle/octave) as baseline where sounds were modulated only in temporal dimension (i.e., they are AM broadband noises). All sounds were sampled at 16 kHz, and root-mean-square (RMS) equalized in power. The stimuli duration was 1 sec with a 10 ms linear ramp at sound onset and offset.

2.3 Procedure

After completing the consent form and questionnaires with personal information, participants were screened by the audiometric testing for twice.

We used an adaptive two-interval, two-alternative forced-choice (2AFC) procedure, and determined the sequence of modulation depth according to the two-down, one-up procedure, which approximates the modulation depth leading to 70.7% positive responses (Levitt, 1971). Specifically, each trial consisted of two sounds, one reference and one target, separated by a 500 ms ISI and randomly presented with equal *a priori* probability. Participants indicated whether the ripples were same or different and the next trial started after the response. Within one block, the spectrotemporal modulations were held constant while the modulation depth varied throughout the procedure, thus resulting in 15 blocks (5 temporal rates multiplied by 3 spectral rates). The initial modulation depth was 40%; and the initial step size was 6% and halved after three reversals. One block ended, either as all 50 trials were finished or as the sequential modulation depth would be beyond the range (1-100%). The order of blocks was randomized and different among participants. An exemplary response pattern was illustrated in Figure 2.

In addition, all participants completed three different reading measures, that is, a one-minute word reading test (Eén-minuut test or EMT, Vorm B; Brus & Voeten, 1973) and two phonological awareness tasks, namely, Gletschr spoonerisms and omkeren (Depessemier & Andries, 2009). The EMT is a test where participants have to correctly read words aloud as fast as possible. Gletschr spoonerisms is a task where participants exchange the onset phonemes of two words (e.g., after hearing "*ruime kabine*" participants should say "*kuime rabine*"). Gletschr omkeren is a task where subjects determine whether the phonemes of the word are correctly reversed (e.g., after hearing "*vats*" and "*stav*" participants should respond "*ja*").

The total duration of the experiment was ~100 minutes, including the psychophysics (five min per block). To avoid the possible fatigue of participants, we divided the psychophysical study into three sessions, each of five blocks, and before each session there was one of the reading measures.

The experiment was conducted in the speech lab (*de Spraaklab*) at Maastricht University. The psychophysics was done in the sound-proof chamber. The auditory stimuli were presented via a standard headphone (SONY MDR-7509HD), and the volume was set at the comfortable level (60 in the system).

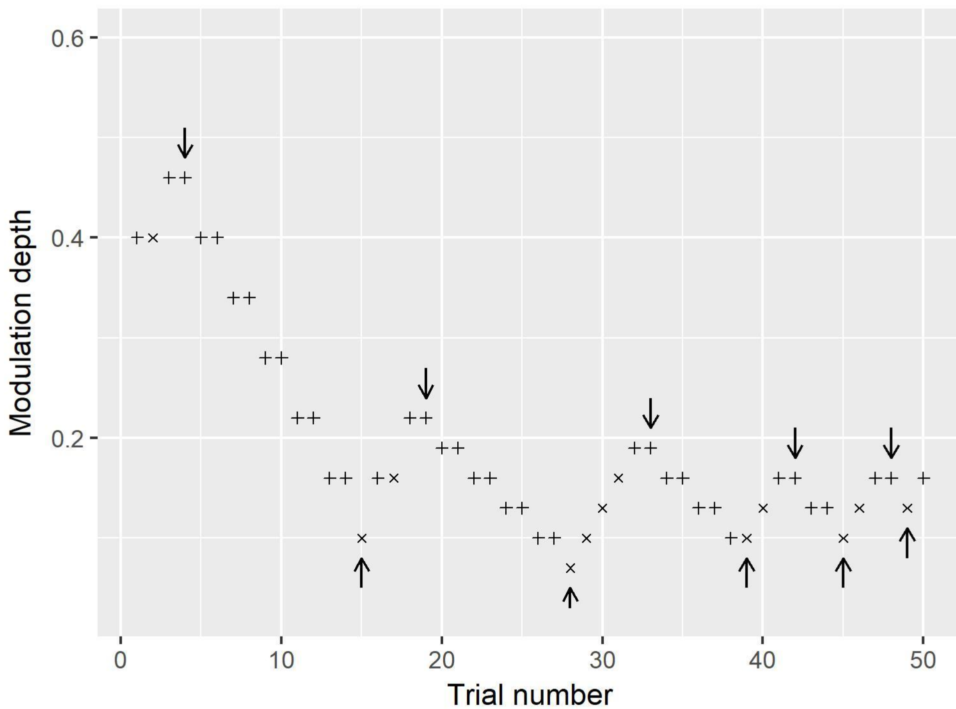

Figure 2. The responses in one block of an exemplary subject. "+" refers to a positive response ("Different"), and "×" negative ("Same"). The modulation depth decreases by one step after two consecutive positive responses and increases following one negative. The reversal, precisely the last trial before change in direction, is denoted by the arrow.

2.4 Data analyses

2.4.1 Reading scores. We calculated standardized reading scores in the following manner: for the EMT, we counted the number of words which were correctly produced; for Gletschr spoonerisms and omkeren, the score was the time needed for completion (in seconds) divided by the number of correct items, then multiplied by 10 (Depessemier & Andries, 2009). Thus, a higher score in the EMT and a lower score in Gletschr spoonerisms and omkeren indicated a better reading performance.

2.4.2 Estimation of detection threshold. We estimated the minimum detection threshold in the same method as described in Levitt (1971) and Chi et al. (1999). That is, it is the mean modulation depth across the last even number of reversals, excluding the first three reversals where the step size was not halved yet. In this way, we could robustly estimate the detection threshold devoid of sampling bias. Most of our data was analyzed in this manner.

However, this method requires a minimum number of five reversals, whereas some (31) blocks consisted of fewer reversals (see Table 1). Specifically, some participants indicated for nearly all trials that they were different (i.e. they gave many positive responses). Consequently, in these blocks, participants reached the minimum modulation depth of the ripples tested in the current experiment (1%) and thus the block stopped before the participants made five reversals in their responses. We argue that it was not due to that participants recognized the experimental procedure or experienced fatigue. First, they showed broad consistency among conditions; and second, the short block did not necessarily appear in the last few ones. For these blocks, we estimated the threshold in the following way: for those with no reversals, the threshold was 0%; and for the others, it was the average of the depths across the reversals and 0%. Thus, the detection threshold would be reliably evaluated.

Table 1. Summary of short blocks (of fewer than 5 reversals) across participants. "D" and "N" refer to dyslexic/normal readers respectively. "T" and "S" refer to temporal/spectral modulations respectively, thus "T2_S1" is the condition with ripples of 2 Hz and 1 cycle/octave. The number between brackets is the order of the occurring block.

Subject	0 reversal	1 reversal	2 reversals	4 reversals	Total
D01			T32_S0 (4) T8_S2.5 (7) T4_S1 (10) T4_S0 (11)	T2_S1 (13)	T2_S1 T4_S0 T4_S1 T8_S2.5 T32_S0
D02	T2_S1 (14)				T2_S1
D05	T4_S0 (4) T16_S0 (8) T4_S2.5 (12) T32_S2.5 (13)	T2_S1 (5)			T2_S1 T4_S0 T4_S2.5 T16_S0 T32_S2.5
D06			T8_S0 (8) T4_S1 (12)		T4_S1 T8_S0
D08	T4_S0 (11) T4_S1 (15)		T2_S0 (9)		T2_S0 T4_S0 T4_S1
D10			T2_S1 (2) T16_S2.5 (5) T8_S1 (13)		T2_S1 T8_S1 T16_S2.5
N01	T16_S1 (11)		T2_S2.5 (12) T4_S2.5 (15)		T2_S2.5 T4_S2.5 T16_S1
N02	T8_S1 (7) T2_S1 (13) T32_S1 (15)		T32_S0 (10) T8_S2.5 (12)		T2_S1 T8_S1 T8_S2.5 T32_S0

					T32_S1
N03			T2_S2.5 (6)		T2_S2.5
N07			T32_S2.5 (12) T4_S2.5 (15)		T4_S2.5 T32_S2.5
N09			T8_S2.5 (11)		T8_S2.5

2.4.3 Statistical testing. To evaluate whether dyslexic and normal readers significantly differ in reading performance, we conducted two-sample two-tailed t tests on scores of the EMT, Gletschr spoonerisms and omkeren. Welch's correction for degree of freedom was used in case that the assumption of equal variance was violated.

We used multilevel modelling (or linear mixed effects modelling) to analyze the psychophysical data. Predictors were selected based on our hypotheses and evaluated in a backwards stepwise manner. Specifically, first, we fitted an omnibus model with Temporal modulations, Spectral modulations, Group and all associated two-way or three-way interactions as fixed effects and a random intercept across subjects. To test whether the three-way interaction showed a significant effect in the first place, we compared the omnibus model with the model excluding this interaction. In case of significance, this predictor was kept, otherwise deleted. Then, we examined the two-way interactions in a similar way. We evaluated the main effects only when the interaction was not significant and thus excluded from the model. For model comparisons, we used chi-square tests to check if the predictor of interest explained significantly more variance in our data. Finally, we conducted post-hoc pairwise t tests with Holm (1979)-correction on *p*-value for multiple comparisons to break down significant interaction effects.

Furthermore, considering that reading scores may explain subtle variance in modulation sensitivity better than the Group predictor, we also fitted the models with EMT, Spoonerisms and Omkeren scores as fixed effects, separately, instead of the Group. We conducted similar backwards fitting as described previously.

All statistical analyses were done using R language (R Core Team, 2012), and relevant packages lme4 (Bates, Mächler, Bolker, & Walker, 2014) and multcomp (Hothorn, Bretz, & Westfall, 2008).

3. Results

3.1 Reading measures

The reading scores (means and standard deviations) of dyslexic and normal readers are presented in Table 2. Two-sample t tests revealed that in all reading measures, dyslexic readers performed significantly more poorly than normal readers, as should be expected. Specifically, for EMT, Welch-corrected $t(21.25) = -4.90$, $p < .001$; for Gletschr spoonerisms, Welch-corrected $t(10.28) = 3.92$, $p < .01$; and for Gletschr omkeren, Welch-corrected $t(13.23) = 6.23$, $p < .001$.

Table 2. Group comparison (group means and standard deviations) in reading performance. ***$p < .001$, **$p < .01$.

	EMT	Spoonerisms	Omkeren
Dyslexic	76.50 (9.57)	125.11 (53.58)	79.64 (14.36)
Normal	96.87 (11.01) ***	56.37 (17.40) **	48.25 (8.50) ***

3.2 Spectrotemporal modulation sensitivity

The estimated detection thresholds (and the standard errors) of dyslexic and normal readers are plotted in Figure 3. Our estimation is broadly consistent with previous studies. Specifically, the tMTF generally resembles a low-pass filter (in particular in spectral rate at 2.5 cycles/octave), consistent with Chi et al. (1999). Moreover, although there are not enough spectral levels to derive a sMTF, the thresholds generally increase as the spectral modulation rate increases, in line with the low-pass sMTF estimated by Chi et al. (1999) and Green (1986). The pure tMTF (i.e., spectral modulation rate = 0) resembles a band-pass filter, inconsistent with Chi et al. (1999). This inconsistency may be the result of a difference in methodology between our studies. That is, Chi et al. (1999) did not test with stationary ripples. Instead, they derived the tMTF only by matrix decomposition of psychophysical data. We argue that the average effects of spectral modulations (i.e., the low-pass sMTF) may introduce bias into the derived tMTF, resulting in a low-pass shape. Moreover, it should be noted that our pure tMTF is in line with the finding of Drullman et al. (1994). Specifically, by removing temporal modulations in speech, they found that modulation rates from 4 to 16 Hz are particularly essential for speech comprehension. Thus, our estimated pure tMTF shares the band-pass shape as the established tMTF for speech intelligibility. In addition, we can observe that the detection thresholds are broadly comparable between dyslexic and normal readers.

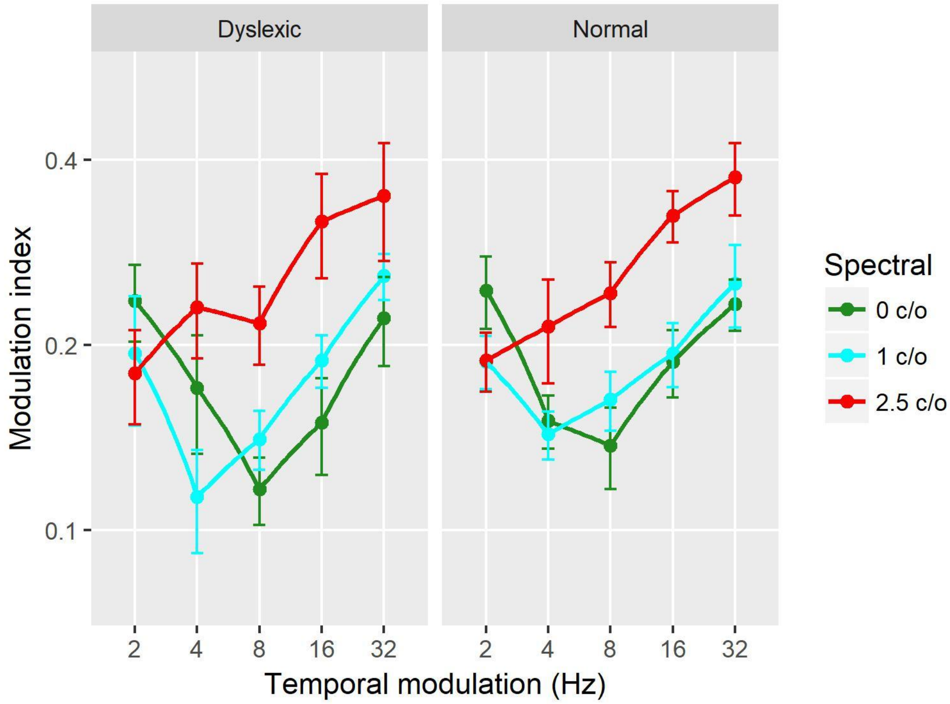

Figure 3. Estimated detection thresholds of dyslexic and normal readers. The error bars represent one standard error of the mean.

A more detailed pattern is illustrated by histograms in Figure 4. Generally, in the conditions with lower detection thresholds, the distribution of detection thresholds is more concentrated and approximates a normal distribution (see “T4_S1” and “T8_S0” for instance), whereas in the conditions with reduced modulation sensitivity, the thresholds do not only increase on average, but the distribution becomes more sparse and covers a broader range (see “T32_S2.5”).

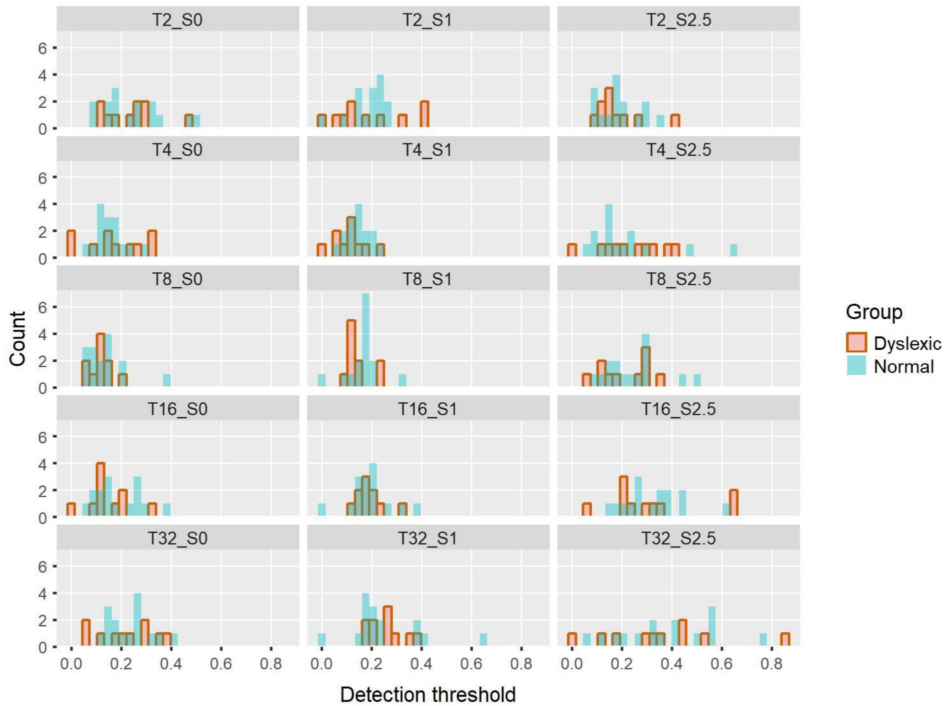

Figure 4. Histograms of detection thresholds of dyslexic and normal readers across conditions.

As depicted in Table 3, model comparisons revealed that there was no significant three-way interaction of Temporal/Spectral modulations and Group (see Table 3, Step 1). Moreover, there was a significant interaction between Temporal and Spectral modulations (Table 3, Step 2, Test 2 vs. 5). In addition, the main effect of Group failed to reach significance level (Table 3, Step 3), whereas there were significant main effects of Temporal and Spectral modulations (Table 3, Step 4, Test 8 vs. 9 and 8 vs. 10).

Model comparisons including reading scores as predictors (Tables 4-6) revealed a similar pattern. That is, the interaction between Temporal/Spectral modulations explained a significant amount of variance in the detection thresholds (Tables 4-6, Step 2, Test 2 vs. 5), yet the effects involving reading scores did not (Tables 4-6, Step 3).

Furthermore, we conducted post-hoc pairwise t tests (Holm-corrected) to break down the significant spectrotemporal interactions. The tests confirmed that the pure tMTF has a band-pass shape while the tMTF for higher spectral modulation rates (2.5 cycles/octave) has a low-pass shape. Specifically, as illustrated in Figure 5, in the pure tMTF, the modulation sensitivity peaks about 8 Hz and is significantly sharper than that for 2 and 32 Hz ($p < .001$); for spectral modulations at 1 cycle/octave, the thresholds for 4 and 8 Hz are significantly lower than for 32 Hz ($p < .01$); and for 2.5 cycles/octave, the thresholds are at the lowest around 2 Hz and increase at 16 and 32 Hz ($p < .01$). In addition, the tests confirmed a reduced sensitivity with increased spectral modulations. Specifically, for temporal modulations at 8 and 16 Hz, the thresholds are significant higher for 2.5 cycles/octave than for 0 and 1 cycle/octave ($p < .01$). For temporal modulations at 32 Hz, there was a trend that the thresholds are higher for 2.5 cycles/octave than for 0 cycle/octave ($p < .1$).

Table 3. Model comparisons with Temporal/Spectral modulations and Group as predictors. “T” stands for Temporal modulations, “S” Spectral modulations and “G” Group. The semi-colon refers to interaction. ***$p < .001$.

Model	Index	df	Log-likelihood	Test	Chi-square	df	*p*
Step 1: to test three-way interaction							
Omnibus	1	32	350.08				
-T:S:G	2	24	348.97	1 vs. 2	2.21	8	.97
Step 2: to test two-way interactions							
-T:S:G -T:G	3	20	348.53	2 vs. 3	0.89	4	.93
-T:S:G -S:G	4	22	348.96	2 vs. 4	0.02	2	.99
-T:S:G -T:S	5	16	327.37	2 vs. 5	43.20	8	<.001***
Step 3: to test the main effect of Group							
T+S+T:S+G	6	18	348.52				
T+S+T:S	7	17	348.45	6 vs. 7	0.14	1	.71
Step 4: to test the main effects of Temporal and Spectral							
T+S	8	9	326.90				
T	9	7	299.36	8 vs. 9	55.08	2	<.001***
S	10	5	293.78	8 vs. 10	66.25	4	<.001***

Table 4. Model comparisons with Temporal/Spectral modulations and EMT scores as predictors. “T” stands for Temporal modulations, “S” Spectral modulations and “E” EMT. The semi-colon refers to interaction. ***$p < .001$.

Model	Index	df	Log-likelihood	Test	Chi-square	df	*p*
Step 1: to test three-way interaction							
Omnibus	1	32	357.90				
-T:S:E	2	24	353.96	1 vs. 2	7.88	8	.44
Step 2: to test two-way interactions							
-T:S:E -T:E	3	20	351.52	2 vs. 3	4.88	4	.30
-T:S:E -S:E	4	22	352.03	2 vs. 4	3.86	2	.14
-T:S:E -T:S	5	16	331.91	2 vs. 5	44.11	8	<.001***
Step 3: to test the main effect of EMT							
T+S+T:S+E	6	18	349.61				
T+S+T:S	7	17	348.45	6 vs. 7	2.34	1	.13

Table 5. Model comparisons with Temporal/Spectral modulations and Spoonerisms scores as predictors. "T" stands for Temporal modulations, "S" Spectral modulations and "Sp" Spoonerisms. The semi-colon refers to interaction. ***$p < .001$.

Model	Index	df	Log-likelihood	Test	Chi-square	df	*p*
Step 1: to test three-way interaction							
Omnibus	1	32	350.08				
-T:S:Sp	2	24	349.46	1 vs. 2	6.04	8	.64
Step 2: to test two-way interactions							
-T:S:Sp -T:Sp	3	20	348.97	2 vs. 3	0.98	4	.91
-T:S:Sp -S:Sp	4	22	348.94	2 vs. 4	1.04	2	0.59
-T:S:Sp -T:S	5	16	327.80	2 vs. 5	43.33	8	<.001***
Step 3: to test the main effect of Spoonerisms							
T+S+T:S+Sp	6	18	348.45				
T+S+T:S	7	17	348.45	6 vs. 7	0.00	1	.98

Table 6. Model comparisons with Temporal/Spectral modulations and Omkeren scores as predictors. "T" stands for Temporal modulations, "S" Spectral modulations and "O" Omkeren. The semi-colon refers to interaction. ***$p < .001$.

Model	Index	df	Log-likelihood	Test	Chi-square	df	*p*
Step 1: to test three-way interaction							
Omnibus	1	32	350.02				
-T:S:O	2	24	349.73	1 vs. 2	0.57	8	.99
Step 2: to test two-way interactions							
-T:S:O -T:O	3	20	348.72	2 vs. 3	2.03	4	0.73
-T:S:O -S:O	4	22	349.53	2 vs. 4	0.41	2	0.81
-T:S:O -T:S	5	16	328.05	2 vs. 5	43.37	8	<.001***
Step 3: to test the main effect of Omkeren							
T+S+T:S+O	6	18	348.51				
T+S+T:S	7	17	348.45	6 vs. 7	0.13	1	.71

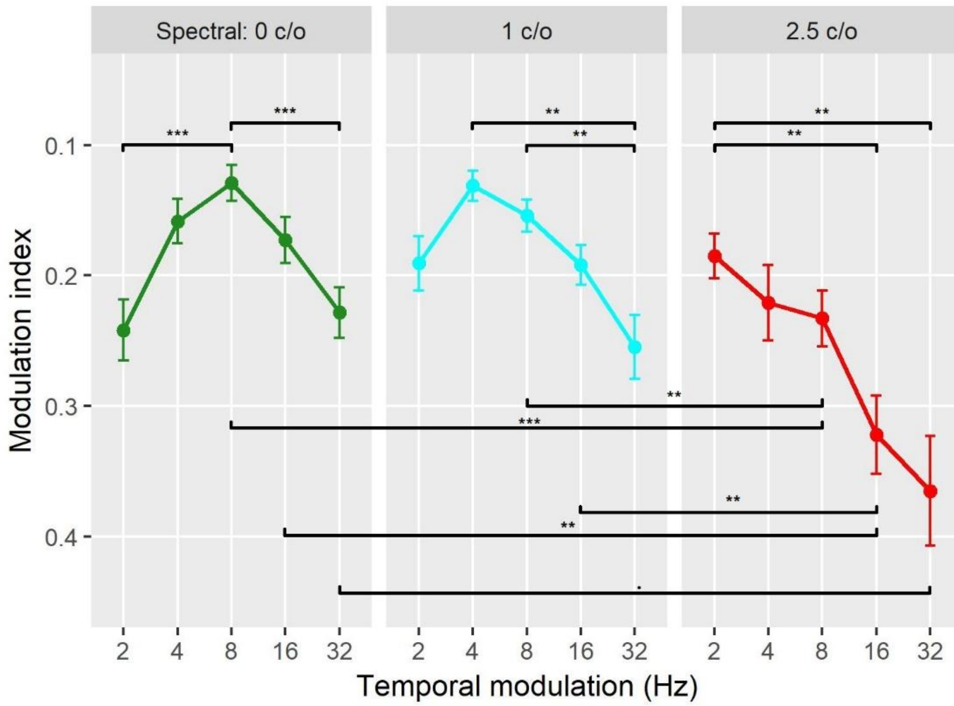

Figure 5. Pooled detection thresholds split by spectral modulations. The error bars represent one standard error of the mean. ***$p < .001$, **$p < .01$, .$p < .1$.

4. Discussion

In sum, we used an adaptive transformed up-down procedure (Chi et al., 1999; Levitt, 1971) to estimate the detection thresholds of dyslexic and normal readers for spectrotemporal modulations in dynamic ripples and AM broadband noises. Contrary to our prediction, there was no significant difference in perceptual sensitivity between dyslexic and normal readers. There were also no significant interactions between the effects of group or reading scores and the effects of spectrotemporal modulations. Whereas added to our expectation, there was a significant interaction between the effects of temporal modulations and those of spectral modulations. The following sections will further address these main findings.

4.1 General auditory deficit hypotheses revisited

As reviewed previously, an extensive number of studies have provided evidence for a general auditory processing deficit underlying developmental dyslexia. Specifically, the amplitude modulation deficit hypothesis (Goswami et al., 2002) proposed a deficiency in perceiving the slow temporal modulations in speech, while the rapid auditory processing deficit hypothesis (Tallal, 1980) argued for a defective processing of rapid spectral transients in speech. Besides these two main theories, some research suggested that the deficit may reside in other aspects of auditory perception such as processing slow spectral modulations (e.g., Talcott et al., 2000). Yet, our results counter all these views. First, we found comparable sensitivity to slow temporal modulations between dyslexic and normal readers, opposing the prediction of the amplitude modulation deficit hypothesis (Goswami et al., 2002). Second, although we did not examine the processing of rapid spectral transients in the current experiment, our results hardly conform with the rapid auditory processing deficit hypothesis (Tallal, 1980). Specifically, dynamic ripples in our study are acoustically more similar to speech than the tone sequence as used in Tallal (1980). That is, we would expect the processing of spectrotemporal modulations resembles speech perception more than does the processing of tone sequence. Hence, it would be hardly conceivable that dyslexic readers perceive tones poorly yet maintain normal spectrotemporal modulation sensitivity. Finally, we found no impairment in perceiving slow spectral modulations, contradicting with the studies such as Talcott et al. (2000). Therefore, as the study of Georgiou et al. (2010), our findings cast doubt on the proposed causal link between general auditory processing deficits and dyslexia.

This null effect cannot be attributed to a small difference in reading level between the two groups, as dyslexic readers performed significantly worse on the reading measures than normal readers. It also cannot be due to an imprecise estimation of detection thresholds. Although the existence of short blocks in our data may induce some estimation error, our results remain consistent with previous psychophysical research on spectrotemporal modulation processing (e.g., Chi et al., 1999).

Ahissar (2007) proposed an alternative hypothesis, the anchoring-deficit hypothesis, to explain the poor processing of rapid auditory stimuli in dyslexia as found in studies such as Tallal (1980). That is, the deficit does not lie in the processing of spectral transients but in the ability to form perceptual anchors. Specifically, dyslexic readers have difficulties with retaining recent sounds in auditory memory to facilitate future processing. This hypothesis is supported by the finding that dyslexic readers discriminate tones normally as the reference tone varies, while as the reference is fixed, normal readers perform significantly better than dyslexic readers (Ahissar, 2007). However, our results seem not to support this theory either. In our case, dyslexic readers displayed normal detection thresholds though the reference was always the same stimulus.

In short, whether a general auditory processing deficit underlies dyslexia remains a debate and is not resolved by our study. Our results join in the other studies which have found negative effects (e.g., Georgiou et al., 2010) and cannot easily reconcile with all present views. If there indeed exists abnormal general auditory processing, we would expect that it impairs the perception of other natural sounds as well. Thus, one way to further evaluate this hypothesis is to investigate the processing of natural sounds in dyslexic readers as opposed to normal readers. Despite that speech is unique and might differ from other complex sounds to some extent, future research following this line will offer additional insight on the relation between general auditory processing and underlying mechanisms of dyslexia.

4.2 Dependency of spectrotemporal modulation processing

As shown previously, we found a significant interaction between the effects of spectrotemporal modulations. Specifically, we discovered that the shape of tMTF changes drastically depending on the spectral modulation rate. That is, the pure tMTF resembles a band-pass filter, while for faster spectral modulations (2.5 cycles/octave), the tMTF has a low-pass shape. This significant interaction implies a dependency between temporal and spectral modulation processing mechanisms.

Previous studies in normal readers have quantified spectrotemporal modulation processing, but the findings remain conflicting on whether these two are independent. For instance, by applying matrix decomposition, Chi et al. (1999) showed that the combined spectrotemporal MTF can be decomposed into a product of temporal and spectral MTFs. Similarly, Langers, Backes and Van Dijk (2003) showed that the brain activation measured with fMRI can be decomposed into the product of fMRI responses to temporal and spectral modulations. These results suggest that the effects of temporal and spectral modulations are separable and thereby indicate independent processing mechanisms. However, other studies found otherwise. For example, Schonwiesner and Zatorre (2009) reported a high degree of interaction between temporal and spectral MTFs estimated from single voxels' fMRI signals. Moreover, Santoro et al. (2014) showed that the joint model employing combined spectrotemporal modulation representations predicts the fMRI responses more accurately than the independent model, suggesting the encoding of joint representations in the brain. Consistently, our data provide further support for the dependency of spectrotemporal modulation processing. That is, the perception of temporal modulations depends on the spectral modulation rate, and vice versa.

One implication of this dependency is that we need to consider the joint effects of spectrotemporal modulations. Thus, the findings obtained with either AM or FM stimuli cannot be safely generalized as a general auditory processing mechanism, since they may miss the complexity of spectrotemporal interaction in the human auditory cortex (Santoro et al., 2014; Schonwiesner & Zatorre, 2009).

Another implication is related to speech processing. Interestingly, speech signals also display a high degree interaction of spectrotemporal modulations. That is, most of the energy in fast spectral modulations is present in slow temporal rates, and the power in fast temporal modulations is limited to slow spectral rates (Elliott & Theunissen, 2009; Singh & Theunissen, 2003). Thus, in addition to the findings that the human auditory system shows enhanced sensitivity to slow spectrotemporal modulations characteristic of speech, the dependency between the effects of spectrotemporal modulations also supports that the human auditory system is optimized for processing the acoustic features of speech. This may reflect an outcome of evolution and/or lifelong behavioral experience where the human auditory system has been shaped by speech (Formisano, 2018).

4.3 Future directions

There are several aspects in our study which can be further investigated in the future. First, we can add a pure spectral modulation condition (i.e., temporal modulation rate = 0) to derive a pure sMTF. It would give us a clearer idea on the sensitivity to spectral modulations in itself. A potential problem in psychophysical studies of spectrotemporal modulation sensitivity is the number of conditions. Thus, to save time and avoid the fatigue effect, we can use a more efficient procedure, such as the weighted up-down method as described in Rammsayer (1992).

Second, we can evaluate the d-prime in dyslexic and normal readers. It will help minimize the possible response bias and thereby increase the validity and/or sensitivity of our estimation.

Finally, we can recruit more participants. Since reading acquisition is a complex cognitive phenomenon and dyslexia is essentially heterogenous, it might be that the subjects we sampled are not impaired in general auditory perception but in other cognition causing reading deficit. Therefore, follow-up research is still needed to inspect this possibility.

References

Abrams, D. A., Nicol, T., Zecker, S., & Kraus, N. (2009). Abnormal cortical processing of the syllable rate of speech in poor readers. *Journal of Neuroscience*, *29*(24), 7686–7693. https://doi.org/10.1523/JNEUROSCI.5242-08.2009

Ahissar, M., Protopapas, A., Reid, M., & Merzenich, M. M. (2000). Auditory processing parallels reading abilities in adults. *Proceedings of the National Academy of Sciences*, *97*(12), 6832–6837. https://doi.org/10.1073/pnas.97.12.6832

Ahissar, Merav. (2007). Dyslexia and the anchoring-deficit hypothesis. *Trends in Cognitive Sciences*, *11*(11), 458–465. https://doi.org/10.1016/j.tics.2007.08.015

Amitay, S., Ahissar, M., & Nelken, I. (2002). Auditory processing deficits in reading disabled adults. *Journal of the Association for Research in Otolaryngology*, *3*(3), 302–320. https://doi.org/10.1007/s101620010093

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406*.5823.

Boets, B., Vandermosten, M., Poelmans, H., Luts, H., Wouters, J., & Ghesquière, P. (2011). Preschool impairments in auditory processing and speech perception uniquely predict future reading problems. *Research in Developmental Disabilities*, *32*(2), 560–570. https://doi.org/10.1016/j.ridd.2010.12.020

Brus, B. T., & Voeten, M. J. M. (1973). Een-minuuttest, vorm A en B, verantwoording en handleiding. *Nijmegen: Berkhout*.

Chi, T., Gao, Y., Guyton, M. C., Ru, P., & Shamma, S. A. (1999). Spectro-temporal modulation transfer functions and speech intelligibility. *The Journal of the Acoustical Society of America*, *106*(5), 2719–2732.

Depessemier, P., & Andries, C. (2009). *GL&SCHR: dyslexie bij+ 16-jarigen: test voor Gevorderd Lezen & SCHRijven*. Garant.

Drullman, R., Festen, J. M., & Plomp, R. (1994). Effect of temporal envelope smearing on speech reception. *The Journal of the Acoustical Society of America*, *95*(2), 1053–1064.

Eddins, D. A., & Bero, E. M. (2007). Spectral modulation detection as a function of modulation frequency, carrier bandwidth, and carrier frequency region. *The Journal of the Acoustical Society of America*, *121*(1), 363–372. https://doi.org/10.1121/1.2382347

Elbro, C., Borstrøm, I., & Petersen, D. K. (1998). Predicting dyslexia from kindergarten: the importance of distinctness of phonological representations of lexical items. *Reading Research Quarterly*, *33*(1), 36–60. https://doi.org/10.1598/RRQ.33.1.3

Elliott, T. M., & Theunissen, F. E. (2009). The modulation transfer function for speech intelligibility. *PLoS Computational Biology*, *5*(3), e1000302. https://doi.org/10.1371/journal.pcbi.1000302

Farmer, M. E., & Klein, R. M. (1995). The evidence for a temporal processing deficit linked to dyslexia: a review. *Psychonomic Bulletin & Review*, *2*(4), 460–493. https://doi.org/10.3758/BF03210983

Formisano, E. (2018). *The "speech ready" auditory cortex*. Manuscript submitted for publication.

Georgiou, G. K., Protopapas, A., Papadopoulos, T. C., Skaloumbakas, C., & Parrila, R. (2010). Auditory temporal processing and dyslexia in an orthographically consistent language. *Cortex*, *46*(10), 1330–1344. https://doi.org/10.1016/j.cortex.2010.06.006

Giraud, A.-L., Lorenzi, C., Ashburner, J., Wable, J., Johnsrude, I., Frackowiak, R., & Kleinschmidt, A. (2000). Representation of the temporal envelope of sounds in the human brain. *Journal of Neurophysiology*, *84*(3), 1588–1598. https://doi.org/10.1152/jn.2000.84.3.1588

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517. https://doi.org/10.1038/nn.3063

Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S., & Scott, S. K. (2002). Amplitude envelope onsets and developmental dyslexia: a new hypothesis. *Proceedings of the National Academy of Sciences*, *99*(16), 10911–10916. https://doi.org/10.1073/pnas.122368599

Goswami, Usha. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, *15*(1), 3–10. https://doi.org/10.1016/j.tics.2010.10.001

Goswami, Usha. (2015). Sensory theories of developmental dyslexia: three challenges for research. *Nature Reviews Neuroscience*, *16*(1), 43–54. https://doi.org/10.1038/nrn3836

Goswami, Usha. (2018). A neural basis for phonological awareness? An oscillatory temporal-sampling perspective. *Current Directions in Psychological Science*, *27*(1), 56–63. https://doi.org/10.1177/0963721417727520

Goswami, Usha, Fosker, T., Huss, M., Mead, N., & Szűcs, D. (2011). Rise time and formant transition duration in the discrimination of speech sounds: the ba-wa distinction in developmental dyslexia: auditory discrimination in dyslexia. *Developmental Science*, *14*(1), 34–43. https://doi.org/10.1111/j.1467-7687.2010.00955.x

Green, D. M. (1986). 'Frequency'and the detection of spectral shape change. In *Auditory frequency selectivity* (pp. 351-359). Springer, Boston, MA.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, 65-70.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical journal*, *50*(3), 346-363.

Kösem, A., & van Wassenhove, V. (2017). Distinct contributions of low- and high-frequency neural oscillations to speech comprehension. *Language, Cognition and Neuroscience*, *32*(5), 536–544. https://doi.org/10.1080/23273798.2016.1238495

Langers, D. R. ., Backes, W. H., & Van Dijk, P. (2003). Spectrotemporal features of the auditory cortex: the activation in response to dynamic ripples. *NeuroImage*, *20*(1), 265–275. https://doi.org/10.1016/S1053-8119(03)00258-1

Leong, V., & Goswami, U. (2014). Assessment of rhythmic entrainment at multiple timescales in dyslexia: evidence for disruption to syllable timing. *Hearing Research*, *308*, 141–161. https://doi.org/10.1016/j.heares.2013.07.015

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. https://doi.org/10.1016/j.jml.2010.09.003

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, *49*(2B), 467–477.

Liu, C., & Eddins, D. A. (2008). Effects of spectral modulation filtering on vowel identification. *The Journal of the Acoustical Society of America*, *124*(3), 1704–1715. https://doi.org/10.1121/1.2956468

Lorenzi, C., Dumont, A., & Füllgrabe, C. (2000). Use of temporal envelope cues by children with developmental dyslexia. *Journal of Speech Language and Hearing Research*, *43*(6), 1367–1379. https://doi.org/10.1044/jslhr.4306.1367

Lundberg, I., Olofsson, åKe, & Wall, S. (1980). Reading and spelling skills in the first school years predicted from phonemic awareness skills in kindergarten. *Scandinavian Journal of Psychology*, *21*(1), 159–173. https://doi.org/10.1111/j.1467-9450.1980.tb00356.x

Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, *54*(6), 1001–1010. https://doi.org/10.1016/j.neuron.2007.06.004

Lyon, G. R. (1995). Toward a definition of dyslexia. *Annals of dyslexia*, *45*(1), 1-27.

McAnally, K. I., Hansen, P. C., Cornelissen, P. L., & Stein, J. F. (1997). The effect of time and frequency manipulation on syllable perception in developmental dyslexics. *Journal of Speech, Language, and Hearing Research*, *40*(4), 912–924.

Menell, P., McAnally, K. I., & Stein, J. F. (1999). Psychophysical sensitivity and physiological response to amplitude modulation in adult dyslexic listeners. *Journal of Speech Language and Hearing Research*, *42*(4), 797–803. https://doi.org/10.1044/jslhr.4204.797

Nicolson, R. I., Fawcett, A. J., & Dean, P. (2001). Developmental dyslexia: the cerebellar deficit hypothesis. *Trends in Neurosciences*, *24*(9), 508–511. https://doi.org/10.1016/S0166-2236(00)01896-8

Overath, T., Zhang, Y., Sanes, D. H., & Poeppel, D. (2012). Sensitivity to temporal modulation rate and spectral bandwidth in the human auditory system: fmri evidence. *Journal of Neurophysiology*, *107*(8), 2042–2056. https://doi.org/10.1152/jn.00308.2011

Pasley, B. N., David, S. V., Mesgarani, N., Flinker, A., Shamma, S. A., Crone, N. E., … Chang, E. F. (2012). Reconstructing speech from human auditory cortex. *PLoS Biology*, *10*(1), e1001251. https://doi.org/10.1371/journal.pbio.1001251

Rammsayer, T. H. (1992). An experimental comparison of the weighted up-down method and the transformed up-down method. *Bulletin of the Psychonomic Society*, *30*(5), 425-427.

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., & Frith, U. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain*, *126*(4), 841–865. https://doi.org/10.1093/brain/awg076

Reed, M. A. (1989). Speech perception and the discrimination of brief auditory cues in reading disabled children. *Journal of Experimental Child Psychology*, *48*(2), 270–292. https://doi.org/10.1016/0022-0965(89)90006-4

Santoro, R., Moerel, M., De Martino, F., Goebel, R., Ugurbil, K., Yacoub, E., & Formisano, E. (2014). Encoding of natural sounds at multiple spectral and temporal resolutions in the human auditory cortex. *PLoS Computational Biology*, *10*(1), e1003412. https://doi.org/10.1371/journal.pcbi.1003412

Santoro, R., Moerel, M., De Martino, F., Valente, G., Ugurbil, K., Yacoub, E., & Formisano, E. (2017). Reconstructing the spectrotemporal modulations of real-life sounds from fmri response patterns. *Proceedings of the National Academy of Sciences*, *114*(18), 4799–4804. https://doi.org/10.1073/pnas.1617622114

Schonwiesner, M., & Zatorre, R. J. (2009). Spectro-temporal modulation transfer function of single voxels in the human auditory cortex measured with high-resolution fmri. *Proceedings of the National Academy of Sciences*, *106*(34), 14611–14616. https://doi.org/10.1073/pnas.0907682106

Singh, N. C., & Theunissen, F. E. (2003). Modulation spectra of natural sounds and ethological theories of auditory processing. *The Journal of the Acoustical Society of America*, *114*(6), 3394–3411. https://doi.org/10.1121/1.1624067

Snowling, M., Nation, K., Moxham, P., Gallagher, A., & Frith, U. (1997). Phonological processing skills of dyslexic students in higher education: a preliminary report. *Journal of Research in Reading*, *20*(1), 31–41. https://doi.org/10.1111/1467-9817.00018

Stanovich, K. E. (1988). Explaining the differences between the dyslexic and the garden-variety poor reader: the phonological-core variable-difference model. *Journal of Learning Disabilities*, *21*(10), 590–604. https://doi.org/10.1177/002221948802101003

Stefanics, G., Fosker, T., Huss, M., Mead, N., Szucs, D., & Goswami, U. (2011). Auditory sensory deficits in developmental dyslexia: a longitudinal erp study. *NeuroImage*, *57*(3), 723–732. https://doi.org/10.1016/j.neuroimage.2011.04.005

Stein, J., & Walsh, V. (1997). To see but not to read; the magnocellular theory of dyslexia. *Trends in Neurosciences*, *20*(4), 147–152. https://doi.org/10.1016/S0166-2236(96)01005-3

Talcott, J. B., Witton, C., McLean, M. F., Hansen, P. C., Rees, A., Green, G. G. R., & Stein, J. F. (2000). Dynamic sensory sensitivity and children's word decoding skills. *Proceedings of the National Academy of Sciences*, *97*(6), 2952–2957. https://doi.org/10.1073/pnas.040546597

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language*, *9*(2), 182–198. https://doi.org/10.1016/0093-934X(80)90139-X

Trehub, S. E., & Henderson, J. L. (1996). Temporal resolution in infancy and subsequent language development. *Journal of Speech Language and Hearing Research*, *39*(6), 1315–1320. https://doi.org/10.1044/jshr.3906.1315

Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. *The Journal of the Acoustical Society of America*, *66*(5), 1364–1380. https://doi.org/10.1121/1.383531

Watson, M., Stewart, M., Krause, K., & Rastatter, M. (1990). Identification of time-compressed sentential stimuli by good vs poor readers. *Perceptual and Motor Skills*, *71*(1), 107–114.