Citation/Reference	Tilde Van Hirtum, Arturo Moncada-Torres, Pol Ghesquière and Jan Wouters **Speech Envelope Enhancement Instantaneously Effaces Atypical Speech Perception in Dyslexia** Ear and Hearing, March 2019
Archived version	Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher
DOI	10.1097/AUD.0000000000000706
Journal homepage	Ear and Hearing
Author contact	tilde.vanhirtum@kuleuven.be; jan.wouters@med.kuleuven.be

Speech Envelope Enhancement Instantaneously Effaces Atypical Speech Perception in Dyslexia

Tilde Van Hirtum[1,2], Arturo Moncada-Torres[1], Pol Ghesquière[2] and Jan Wouters[1]

[1]KU Leuven – University of Leuven,
Department of Neurosciences, Research group Experimental ORL
O&N 2, Herestraat 49 – box 721
B-3000 Leuven
Belgium

[2]KU Leuven – University of Leuven,
Faculty of Psychology and Education Sciences, Parenting and Special Education Research Unit
Leopold Vanderkelenstraat 32 - box 3765
B-3000 Leuven
Belgium

Corresponding author
Tilde Van Hirtum
O&N 2, Herestraat 49 – box 721
B-3000 Leuven
+32 16 37 43 55
tilde.vanhirtum@kuleuven.be

Conflict of Interest and Sources of funding
There are no conflicts of interest to disclose.

This work was funded by the Research Foundation - Flanders (FWO) (G.0A91.15N).

ABSTRACT

Objectives: Increasing evidence exists that poor speech perception abilities precede the phonological deficits typically observed in dyslexia, a developmental disorder in learning to read. Impaired processing of dynamic features of speech, such as slow amplitude fluctuations and transient acoustic cues, disrupts effortless tracking of the speech envelope and constrains the development of adequate phonological skills. In this study, a speech envelope enhancement (EE) strategy was implemented to reduce speech perception deficits by students with dyslexia. The EE emphasizes onset cues and reinforces the temporal structure of the speech envelope specifically.

Design: Speech perception was assessed in 42 students with and without dyslexia using a sentence repetition task in a speech-weighted background noise. Both natural and vocoded speech was used to assess the contribution of the temporal envelope on the speech perception deficit. Their envelope-enhanced counterparts were added to each baseline condition to administer the effect of the EE-algorithm. In addition to speech-in-noise perception, general cognitive abilities were assessed.

Results: Results demonstrated that students with dyslexia not only benefit from envelope enhancement, but benefit more from it than typical readers. Hence, envelope enhancement completely normalized speech reception thresholds for students with dyslexia in adverse listening conditions. In addition, a correlation between speech perception deficits and phonological processing was found for students with dyslexia, further supporting the relation between speech perception abilities and reading skills. Similar results and relations were found for conditions with

natural and vocoded speech, providing evidence that speech perception deficits in dyslexia stem from difficulties in processing the temporal envelope.

Conclusions: Using speech envelope enhancement speech perception skills in students with dyslexia were improved passively and instantaneously, without requiring any explicit learning. Additionally, the observed positive relationship between speech processing and advanced phonological skills, opens new avenues for specific intervention strategies that directly target the potential core deficit in dyslexia.

INTRODUCTION

After decades of dyslexia research, the prevailing view of dyslexia connects deviant phonological processing (i.e. the ability to process the sounds of one's language and their mental representation) with the severe and persistent reading difficulties universally observed in children and adults with dyslexia (Goswami, 2000; Snowling, 2000). Recent theories have considered that subtle speech perception deficits might give rise to these phonological problems (Ziegler et al. 2009; Boets et al. 2011; Goswami 2015; Vanvooren et al. 2017). However, speech includes a great deal of redundant information and in optimized listening conditions (nearly) perfect speech perception can be achieved despite a considerable amount of spectral or temporal degradation (Shannon et al. 1995). Consequently speech perception deficits might be obscured in quiet conditions, leading to rather mixed results in past research (Brady et al. 1983; Joanisse et al. 2000; Hazan et al. 2009; Ziegler et al. 2009; Robertson et al. 2009; Messaoud-Galusi et al. 2011; Calcus et al. 2016). Conversely, the postulated speech perception deficits emerge clearly in a background of noise, and so results from testing speech perception in noise can be used to predict phonological and literacy skills in children and adults with dyslexia (Brady et al. 1983; Ziegler et al. 2009; Boets et al. 2011; Poelmans et al. 2011; Dole et al. 2012; Calcus et al. 2016). It is therefore hypothesized that people with dyslexia cannot adequately extract important acoustic information of the speech stream, which becomes even more difficult in noisy environments. This general lack of speech robustness hinders the encoding of stressed syllables, syllables, and onset time, and complicates the establishment of stable phoneme categories. Consequently, this results in difficulties with mapping letters (graphemes) and their corresponding speech sound representations (phonemes) when children are learning to read (Ziegler et al. 2009).

Speech signals contain a broad variety of acoustical cues. It is important to determine precisely which aspects of the speech signal constrain sensory processing in dyslexia the most, in order to be able to improve speech perception. Recent evidence suggests that mostly impaired processing of dynamic features of speech at syllable-relevant rates, such as overall amplitude changes (amplitude modulation) and transient acoustic cues (onset time), disrupts effortless tracking of the speech envelope and constrains the development of adequate phonological skills (Goswami 2011). Numerous studies over the last decades supported this hypothesis, suggesting that the ability to process the dynamics of the speech envelope accurately is the underlying core deficit in dyslexia (Menell et al. 1999; Lorenzi et al. 2000; Witton et al. 2002; Goswami et al. 2002; Richardson et al. 2004; Goswami et al. 2010; Leong et al. 2011; Poelmans et al. 2011; Power et al. 2016). Alternatively, vocoded speech signals have been used to increase reliance on speech envelope cues (Shannon et al. 1995; Power et al. 2016). Indeed, previous studies using these vocoded speech signals have shown that children with dyslexia had poorer speech recognition compared to typical readers (Ziegler et al. 2009; Johnson et al. 2011; Nittrouer & Lowenstein 2013; Power et al. 2016), additionally supporting the framework proposed by Goswami (Goswami 2011). However, by using only one type of stimuli (e.g. vocoder) it is impossible to know whether poor speech perception in dyslexia resulted from poor amplitude envelope processing or from poor processing of degraded speech signals more generally. Subsequently, an alternative hypothesis suggests that a general, perceptual noise exclusion deficit hinders speech perception, which in turn could lead to difficulties with phonological awareness and multisensory integration, both of which are linked to reading development (Sperling et al. 2005; Ziegler et al. 2009). This hypothesis implies that deficient noise exclusion rather than abnormal temporal processing affects the acquisition of literacy skills.

Therefore, the first goal of this study was to investigate speech perception in a stationary background noise and in a condition of an additional speech signal degradation. In addition to natural speech, we implemented an eight channel noise-band vocoder. By comparing speech perception of natural speech to vocoded speech, where we limit temporal fine structure as well as spectral information and relocate the primary focus on the temporal envelope (Shannon et al. 1995), we aimed at disentangling both hypotheses. Considering the framework proposed by Goswami (Goswami et al. 2002; Goswami 2011), we expected participants with dyslexia to have difficulties processing the dynamics of the envelope, and we hypothesized that they would perform worse than typical readers when they were exposed to natural speech signals as well as to speech signals with degraded spectral resolution and limited temporal fine structure. However, if speech perception deficits are due to a general lack of noise exclusion we would expect larger group differences for the vocoded signals, which are further degraded.

The second and main goal of this study was to not only establish the presence and the nature of the speech perception deficit in dyslexia, but to improve speech processing in general. Hence, if speech-in-noise difficulties in dyslexia indeed emerge from altered speech envelope tracking (Goswami 2011), then subsequently enhancing the envelope should improve speech perception by persons with dyslexia. Over the last decades, a variety of studies have investigated the benefit of enhancing envelope cues motivated by the idea initially introduced by Schroeder (1981). The original work of Langhans & Strube (1982) showed substantial potential from envelope enhancement processing. However following studies manipulating the whole envelope found rather moderate effects of envelope expansion (Clarkson et al. 1991; Fu & Shannon 1999) or even clear detrimental effects of envelope expansion (Plomp 1988; van Buuren et al. 1999). Therefore,

in this work we focused on emphasizing specific parts of the envelope, transient acoustic cues, which seem to constrain effortless speech processing in dyslexia the most.

We implemented the speech envelope enhancement (EE) strategy developed by Koning & Wouters (2012) to investigate the possibility of improving speech perception skills of students with dyslexia. By using this signal processing algorithm, originally developed for cochlear implant users, we have the substantial benefit of potential real-time implementation, which to our knowledge has not been tested until now. The EE algorithm amplifies amplitude transitions in the speech envelope explicitly, without altering other parts of the signal. We aim to facilitate the detection of important envelope fluctuations in order to strengthen information processing of syllable onsets and phoneme discrimination in dyslexia. We hypothesize that emphasizing these challenging dynamic features might ameliorate speech processing and could potentially restore speech sound representations in our population with dyslexia.

MATERIALS AND METHODS

Participants

Forty-four higher education students participated in this study: 21 with dyslexia (age range from 18 to 25 years, 10 male) and 23 typical readers (age range from 18 to 26 years, 6 male). Participants were recruited through advertising at the University and the Clinical Centre (PraxisP) of Leuven. Participants with dyslexia had received a formal diagnosis of dyslexia independent of the study. All participants with dyslexia showed persistent reading and spelling difficulties by scoring below the 10th percentile on standardized word-reading (Brus & Voeten 1979), pseudo word-reading (Van den Bos et al. 1994) and spelling tests (Ghesquière 1998). All typical readers were required to score above the 20th percentile on the same tests. All participants were native Dutch speakers,

had normal peripheral hearing (hearing thresholds ≤ 25 dB HL for frequencies from 0.5 to 8 kHz), adequate non-verbal intelligence (≥ 85; WAIS-III (Wechsler 1997)) and no other neurological or developmental disorders. Two participants, initially recruited for the typical reading group, did not comply with the recruitment criteria (one with reading scores < 10th percentile and one with dyscalculia) and were therefore excluded for further analyses. The study was approved by the ethics committee of the University of Leuven (approval number B322201526287) and written consent was obtained from each participant before participation.

Cognitive tasks

We designed an extensive set of cognitive tasks to compare general cognitive ability in the two groups. All main phonological skills were tested independently. Firstly, a spoonerisms task, in which participants were asked to exchange the initial letters of two words or pseudo words, was used to assess phonological awareness (PA) (Poelmans et al. 2011). Both task accuracy and response time were taken into consideration. Secondly, phonological fluency was measured through four rapid automized naming (RAN) tasks: objects, colors, digits and letters (Vanvooren et al. 2017). Lastly, the Digit Span subset of the WAIS-III was administered (Wechsler 1997). Participants were required to repeat an increasing series of digits, in forward and reverse order. The digit span forward gave an indication of verbal short-term memory (VSTM) whereas digit span backward (WAIS-III; Wechsler, 1997) explored working memory. In addition, language skills and verbal intelligence were determined through the vocabulary subtest of the WAIS-III (Wechsler 1997). Detailed group comparisons for each of these cognitive tasks can be found in Table 1.

Signal processing

Signal processing was performed using Matlab R2013b (The MathWorks Inc 2013). We tested the participants' speech understanding in noise in four different conditions: natural speech, vocoded speech, and their envelope-enhanced (EE) counterparts (see Fig. 1). The natural speech signal was chosen as a reference.

Vocoded speech was used as a model for simulating reduced access to spectral cues and temporal fine structure. To create the vocoded speech, the signal was first decomposed into eight channels using an analytical filter bank consisting of fourth order Butterworth filters. Cut-off frequencies for the eight channels were 187.5, 437.5, 687.5, 1062.5, 1562.5, 2312.5, 3437.5, 5187.5 and 7937.5 Hz (Koning & Wouters 2012). Then, the envelope of each channel was extracted by filtering the half-wave rectified signal with a fourth-order Butterworth filter with f_c = 200 Hz. This envelope was used to modulate a broadband noise carrier, which was filtered using a synthetic filter bank with the same parameter used in the analysis stage. Finally, all eight noise vocoded channels were level matched to their original in-band input and summed to obtain the final vocoded signal.

The signal processing methods used to generate the EE signals was adapted from Koning and Wouters (Koning & Wouters 2012). First, the input signal was resampled to $f_{resampling}$ = 16 kHz and divided into frames of 128 samples with a frame advance of 32 samples. Then, the signal was transformed into the frequency domain by computing the Fast Fourier Transform using a Hann analysis window. The envelope of each channel $E_{(t,k)}$ was extracted using a weighted summation of the power of its respective frequency bins. The frequency limits for the different channels were defined according to the critical bandwidths of Fastl and Zwicker (Fastl & Zwicker 2007) up to $f > \frac{f_{sampling}}{2}$. Then, $E_{(t,k)}$ was processed to generate a slow envelope $E_{slow(t,k)}$ by using a fourth order Butterworth filter with a cut-off frequency of 20 Hz. This removes F0 and increases the time

delay of $E_{slow(t,k)}$ when reacting to sudden increases in $E_{(t,k)}$. Afterwards, $E_{slow(t,k)}$ was amplified by a factor A_{slow} = 8 to make sure that its level is higher than that of $E_{(t,k)}$ at quasi-stationary parts. At a sudden increase in energy (i.e. an onset), $E_{(t,k)} > E_{slow(t,k)}$. The peak signal $E_{peak(t,k)}$ was extracted by subtracting the amplified $E_{slow(t,k)}$ from $E_{(t,k)}$. Then, half-wave rectification ensured that no negative values (which occur in speech parts with stationary levels) are present. The half-wave rectified $E_{peak(t,k)}$ had values different from zero at onsets of $E_{(t,k)}$. $E_{peak(t,k)}$ was amplified by a factor A_{peak} = 18 to obtain the final $E_{peak(t,k)}$ signal. The enhanced envelope $EE_{(t,k)}$was obtained by adding $E_{(t,k)}$ and $E_{peak(t,k)}$ together. Finally, the time signals $E_{peak(t)}$ and $EE_{(t)}$ were obtained by computing the inverse Fast Fourier Transform using a Hann synthesis window. The average periodicity at which individual peaks appeared in $E_{peak(t)}$ across all sentences was 3.0 Hz, corresponding to the syllable rate of the speech material (Van Wieringen & Wouters 2008). Additionally AMi spectra were calculated as described by Varnet et al. (2017), showing a significantly stronger peak after EE processing between 2.5 and 3.2 Hz. Further details of the EE signal processing are described in Koning and Wouters (Koning & Wouters 2012).

To enhance the signals, $E_{peak(t)}$ was computed from the natural, and unprocessed signal and added to the natural and vocoded signals respectively using broadband summation. Note that, contrary to the method described in Koning and Wouters (Koning & Wouters 2012), here identical envelopes $E_{peak(t)}$ are applied to both input signals to achieve onset enhancements at exactly the same locations. A maximal scaling factor (i.e., for all the speech material; 10.5 dB) was computed to avoid that the enhanced signals exceeded the dynamic range. This way, we guaranteed that the stimuli presented to the participant would make use of the whole available dynamic range.

Enhancing the signals by EE increased their root mean square values (RMS). The perceptual relevance of this increase in RMS was assessed in a preliminary loudness rating experiment, in which the participants had to rate the loudness of randomly selected stimuli (with or without EE). Mean loudness scores did not differ between the four different conditions (all $P > 0.05$).

Speech perception task

We assessed speech understanding using the Leuven Intelligibility Sentence Test (Van Wieringen & Wouters 2008). The full corpus consists of 35 lists, each of 10 Dutch, semantically correct sentences uttered by a female speaker. Each list of 10 sentences contains 32 or 33 keywords that were counted in the computation of the percentage-correct score. The lists of sentences were presented in a stationary speech-weighted (vocoded) noise, matching the long-term average spectrum of all sentences together. In all four conditions, the noise level was fixed at 65 dB SPL and the speech level was adjusted to obtain the targeted signal-to-noise ratio (SNR). Sentences were presented to the participants monaurally to the right ear at four fixed SNRs in 2 dB steps from -7 to -13 dB for natural speech, and +1 to -5 vocoded speech. Additional lists at ± 2 dB SNR were added if this was necessary to obtain an accurate psychometric function. All participants received a short training session with 10 natural speech and 10 vocoded sentences at -5 dB and +5 dB respectively. Every condition and SNR was tested and retested in two sessions for each participant. Each participant listened to 340 sentences in total (= 20 training sentences + 10 sentences/condition x 4 SNRs x test- and retest session). Which list of 10 sentences was presented under each of the four conditions and under each of the four SNRs was randomized across participants. All conditions were presented to the participants in random order to limit learning effects (i.e. a random subgroup performed the vocoded task first) and every list of sentences was only presented once. A

Scarlett 2i2 Focusrite soundcard and DTH005 headphones were used to present the stimuli to the participants using the software platform APEX (Francart et al. 2008).

Statistical analysis

All statistical analyses were done in R (version 3.2.3) (R Core Team 2015). Firstly, to assess test-retest reliability of the individual scores, the within-subject standard deviation (i.e. measurement error) was calculated by dividing the standard deviation of the test-retest differences by the square root of 2 (Smits et al. 2004). No significant differences between test sessions that exceeded the calculated measurement error of 0.72 dB were found, so the scores of each individual for the two sessions were averaged for further analysis. Individual psychometric curves were fitted to the percentage correct scores with a cumulative Gaussian function using a generalized linear model. The speech reception threshold (SRT) was calculated from each individual model as the SNR corresponding to a percentage correct score of 50%. The precision (error bar) on the SRT was determined as the standard error along the x axis (in dB) (MASS package). Average SRTs were calculated as the arithmetic average of all individual SRTs whereas the overall precision on the SRT was computed as the quadratic average of all error bars. The speech material used (Leuven Intelligibility Sentence Test) in combination with the model fitting approach, allowed a very precise determination of the SRT (precision ≤ 0.3 dB, Table 2).

The SRTs between participants with and without dyslexia were compared by means of linear mixed models using maximum likelihood criteria (nlme package); EE was included as a nested random effect by participant, reflecting the repeated-measures predictor. In addition, the interaction between participant group and EE was taken into consideration. Significant interaction terms were further explored by using two-tailed post-hoc Tukey Contrasts ($\alpha = 0.05$). Only adjusted p-values

are reported. Residuals were checked with a Shapiro-Wilk's test for normality and Levene's test for equality of variances. Both assumptions were met at each level of the predictor variables.

A two-step procedure (Ramus et al., 2003) was used to calculate z-scores and identify the participants demonstrating abnormal performance for all cognitive and speech tasks. A control mean and standard deviation were calculated based on the z-scores of the typical readers. Then, the control mean and standard deviation were recomputed without typical readers scoring below the fixed threshold of -1.65xSD, to avoid overestimation of the normal range of performance. Based on the recomputed threshold, dyslexics who performed outside ± 1.65xSD were identified. Z-scores of the typical readers without EE were used as the baseline, to identify individuals with dyslexia with specific speech perception deficits before and after EE-implementation. This baseline reflects the threshold participants with dyslexia should reach with EE to simulate normal speech-in-noise performance.

Composite z-scores were created for literacy (word-reading, pseudo word-reading, and spelling tasks), phonological awareness (spoonerism task; accuracy and response time) and phonological fluency (RAN; objects, colors, digits, and letters). Digit span forward and backward z-scores were not averaged, since these tasks explain different dyslexia-related difficulties and results from them were only weakly correlated. Subsequently, these composites and z-scores were used to explore the relation between speech perception and literacy when controlling for different possible confounding variables. Step-wise regression analyses were carried out to determine the unique variance in literacy skills accounted for by speech perception (R^2 change) while correcting for the potential confounds (see Table 3 & 4).

RESULTS

Speech perception of natural speech

Psychometric functions and individual SRTs for both groups and speech conditions are illustrated in Figure 2 and descriptive data is provided in Table 2.

We observed significant differences in speech processing without EE between participants with dyslexia and typical readers ($\chi^2(1) = 9.66$, $p = 0.0019$). SRTs were on average 0.87 dB worse (i.e. higher SRTs) for participants with dyslexia as determined by post-hoc tests on the linear model (z = -3.7, $p = 0.001$, 95% CI [0.28, 1.45]). After applying EE, we found a significant improvement in the SRT in both groups ($\chi^2(1) = 47.18$, $p < 0.001$). Strikingly, the interaction term between EE and group was also significant ($\chi^2(1) = 4.34$, $p = 0.03$), showing that the benefit derived from EE was greater for the group with dyslexia (z = 2.138, $p = 0.03$, 95% CI [0.03, 0.57]). Moreover, this larger improvement in SRT effectively bridged the gap in speech perception between both groups, so that adults with dyslexia no longer showed any deficiency when EE was applied (z = -2.427, $p = 0.06$, 95% CI [-0.015, 1.16]).

To further strengthen these findings, we explored individual deviance profiles using a two-step procedure, to identify individuals with specific speech perception deficits (Ramus et al., 2003). The majority of the participants with dyslexia (~53%), but only a few of the typical readers (~9%) exhibited specific speech perception deficits These results not only show that speech-in-noise perception difficulties occur in dyslexia, but are the first to show a deficit in more than half of the dyslexic group. In addition, deviance analysis confirmed a significantly greater likelihood of speech perception difficulties in participants with dyslexia ($p = 0.005$, Fisher's exact test, odds ratio 9.84, 95% CI [1.67 108.5]). However, speech perception improved for 95% of the participants in

the dyslexic group when EE was applied, and speech perception deficits could no longer be demonstrated (p = 0.66, Fisher's exact test, odds ratio = 2.19, 95% CI [0.27 27.2]). In conclusion, these results show an approximation of normal speech perception skills using EE for participants with dyslexia at group level as well as at the individual level.

Speech perception of vocoded speech

Overall, average SRTs were 7.2 dB worse than for natural speech. Hence, perception of vocoded speech was considerably more difficult for both participants with dyslexia and typical readers. Nevertheless, we found clear deficits in the perception of vocoded speech, comparable to those obtained with natural speech in noisy conditions. Again, participants with dyslexia showed significantly worse speech processing skills than typical readers when EE was not applied ($\chi^2(1)$ = 8.11, p = 0.004) and SRTs were on average 0.88 dB worse for participants with dyslexia than for typical readers (z = 2.85, p = 0.004, 95% CI [0.27, 1.48]) (see Fig. 3 and Table 2). Furthermore, both groups presented a significant improvement in the SRT with EE ($\chi^2(1)$ = 25.96, p < 0.001). The absolute benefit of the EE appeared to be similar for both groups since the interaction term between EE and group was not significant. Additionally, Fisher's exact test revealed that a significantly higher number of participants with dyslexia had specific speech-in-noise difficulties (p = 0.02, Fisher's exact test, odds ratio = Inf, 95% CI [1.37 Inf]). Around 30% of the participants with dyslexia displayed deviant scores, compared to none of the typical readers. However, after implementation of EE, participants with dyslexia again approximated normal speech perception skills on an individual level, and as a group they no longer demonstrated speech perception deficits (p = 0.23, Fisher's exact test, odds ratio = Inf, 95% CI [0.43 Inf]).

To sum up, participants with dyslexia were not disproportionately hindered by the spectral degradation of the vocoded signal. In addition, individuals that had difficulties with vocoded speech, all had difficulties with the natural speech as well. Together, we provide significant evidence that speech perception deficits in dyslexia mainly stem from difficulties in processing the temporal envelope, in favor of the framework of Goswami (Goswami et al., 2002; Goswami, 2011).

General cognitive abilities and stepwise regression analyses

Finally, we considered the impact of general cognitive abilities on the speech perception differences we found between participants with dyslexia and typical readers, given that intelligence, verbal skills and memory performance are involved in complex speech-in-noise tasks, and may potentially influence performance in both groups (Ramus, 2001). Our results indeed show significantly better scores for typical readers than for participants with dyslexia on the vocabulary task ($p = 0.006$) and all phonological processing tasks including phonological awareness ($p < 0.001$), phonological fluency ($p < 0.001$) and verbal short-term memory ($p < 0.001$) (see Table 1). However, step-wise regression analyses indicated that the SRT is a significant predictor of literacy, even after taking into account these cognitive task variables. The amount of unique variance explained by the SRT ranged between 4 and 29% for natural speech (see Table 3) and between 3 and 14% for vocoded speech (see Table 4). In addition, we observed a clear improvement in SRT after implementing EE, while the task remained the same. Consequently, if group differences in general cognitive abilities would explain the observed speech perception deficits, these deficits should still be present regardless of EE. Taken together, our results provide strong evidence that impoverished speech perception in participants with dyslexia cannot be explained by their lower general cognition, phonological skills or reduced verbal abilities.

Furthermore, we investigated the overlap between phonological processing and speech-in-noise deficits. We observed that all individuals with speech perception deficits had deviant scores on at least one task, 90% for two tasks and 36% for all three tasks related to phonological processing. In addition, we revealed a significant correlation between speech perception and phonological skills ($r = 0.47$, $p = 0.02$) in participants with dyslexia, indicating that those with better speech perception skills had better phonological awareness outcomes. Although a correlation does not imply causality, we believe that these results provide evidence for an explicit association between speech perception and phonological processing in dyslexia in contrast to previous research (Joanisse et al. 2000; Robertson et al. 2009).

DISCUSSION

Sensory impairments in processing speech envelope information arise and affect phonological development starting from birth in dyslexia (Goswami 2011; Goswami et al. 2016). Through repeated misperceiving of specific features of sound (i.e. slow amplitude modulations), faulty sensory representations can develop in children with dyslexia (Ziegler et al. 2009; Hornickel et al. 2012). However, children refine their phonological representations when they begin to learn grapheme-phoneme correspondences (Johnson et al. 2011). Typically this is taught by the teacher in a classroom condition, and by default noisy environment, which constrains sensory processing disproportionately for children with dyslexia. In this study, we implemented a signal processing algorithm to assess whether enhancing these difficult parts of the speech signal (i.e. the onsets) might improve their inadequate speech perception. Here, we demonstrated the feasibility of using an envelope enhancement strategy to instantaneously correct poor speech-in-noise perception in dyslexia.

To assess our participant's abilities to specifically use the temporal envelope in their speech perception, we compared the performance of perception of natural speech with noise-vocoded speech. Using vocoded speech, intelligibility is highly dependent on the number of available spectral channels. Near perfect intelligibility of sentences in quiet can be obtained using as limited as 4 to 6 channels (Shannon et al. 1995; Dorman et al. 1997; Friesen et al. 2001; Shannon et al. 2004). In more difficult listening conditions, such as background noise, more than 6 channels are needed to perform better than 50% correct, however the point at which performance asymptotes in noise depends strongly on the SNR, the characteristics of the noise and the speech material (Shannon et al. 2004; Stilp & Goupell 2015). Taken together, an 8-channel vocoder should provide sufficient spectral information for adult participants without clear sensory impairments or spectral processing disorders. Our results showed that participants with dyslexia were not disproportionately hindered by the degraded spectral resolution and limited temporal fine structure of the noise-vocoded speech. The evidence found here, therefore, suggests a specific temporal processing deficit for these adults with dyslexia rather than a general noise exclusion deficits or spectral processing disorders (see also Lorenzi et al. 2000; Goswami et al. 2002; Ziegler et al. 2009; Goswami 2011; Nittrouer et al. 2018). It may be the case that when fewer than 8 channels are presented, performance might decrease significantly for participants with dyslexia and group differences would become larger for the vocoded signals irrespective of a general noise exclusion deficit. With 8 channels, participants have, although restricted, sufficient spectral information to process speech and potentially compensate partly for the temporal processing difficulties similarly to natural speech perception. However when fewer channels are available, reliance on the envelope will become stronger to the point where spectral compensation will no longer be an option. This suggestion has been made by Nittrouer & Lowenstein (2013) where children with dyslexia had larger group differences for vocoded signals compared to unprocessed speech. However, future

research showing how different numbers of available spectral channels influences speech perception is necessary to further confirm such a hypothesis.

It may seem surprising that the effect of EE seemed smaller for the noise-vocoded condition for the participants with dyslexia considering the increased reliance on envelope cues. It may be the case that the overall effect of EE was smaller due to choosing the same parameter value (A_{peak}) for both speech materials and the adding of the envelope to the signal. This resulted in overall fewer and less pronounced peaks which can be seen clearly in Figure 1. An alternative interpretation might be that for vocoded speech participants with dyslexia need more training with the speech material, and consequently with the algorithm, considering that vocoded speech is significantly more difficult compared to natural speech. It is well known that normal listeners need at least some training to be able to perform well on tasks including vocoded speech (e.g. Koning & Wouters 2012; Rosen et al. 2015). The reduced benefit of EE in de vocoded condition leads us to believe that the potential of EE might be obscured by learning effects for participants with dyslexia. However, the significance of the study lies in the fact that the present findings represent instantaneous effects after only limited exposure to the EE-algorithm, which are present in both conditions. We are willing to hypothesize that after long-term, repeated exposure, EE will have even stronger effects on speech perception in both conditions.

Hence, it was possible to improve speech perception abilities using EE in both speech conditions by helping participants with dyslexia to extract relevant dynamic features from running speech. A possible theoretical explanation of this observation might be found in current speech perception models. Envelope rise times serve as acoustic landmarks or temporal ‘edges’ in the signal, which facilitate temporal segmentation and increase tracking of the speech envelope (Goswami 2011;

Gross et al. 2013; Doelling et al. 2014). A growing number of studies have shown reduced neural tracking in dyslexia at critical frequencies (~2-4 Hz) of the speech envelope (Hämäläinen et al. 2012; Soltész et al. 2013; Power et al. 2016; Molinaro et al. 2016). Our EE algorithm amplifies these underspecified rise times in the envelope approximately at syllable-rate, which coincides with the range of frequencies where atypical neural processing was found in dyslexia. Therefore we hypothesize that the EE strategy enhanced their suboptimal neural mechanism by increasing phase-locking to these acoustic landmarks and strengthening envelope tracking, resulting into the perceptual improvement we observed.

One significant advantage of EE as a potential intervention strategy is the possibility of real-time implementation. The use of assistive listening devices has already been proven to boost neurophysiological responses to speech along with literacy and phonological abilities in children with dyslexia (Hornickel et al. 2012). However, improving the signal-to-noise ratio will increase acoustic clarity, but might not necessarily boost speech envelope tracking nor might it restore the altered speech sound representations. Considering that the EE algorithm was originally developed for assistive listening devices, it might be adjusted efficiently for similar FM-system use. Concern may still exist about the effectiveness of the EE algorithm in real-life noise levels since the tested SNRs fall largely outside the realistic range of daily-life acoustic environments (Pearsons et al. 1977; Smeds et al. 2015). Our task required participants to repeat sentences at SNRs ranging from -7 to -13 dB SNR, and performance of both groups appears to be at ceiling around -5 dB SNR. Hence, not much effect could be measured for natural speech material above -5 dB SNR. In contrast, for vocoded speech a significant improvement could be measured above 1 dB SNR, thus the EE-algorithm proved to be effective not only across a range of at least 14 dB SNR, but additionally at more realistic SNRs in more difficult listening conditions. In addition, fixed SNRs

depend strongly on the speech material (e.g. sentences, words), the background noise (e.g. stationary, fluctuating), but also the characteristics of the participant (e.g. age) (e.g. Francart et al. 2011; Goossens et al. 2017). In the present study we investigated the effect on a group of high performing higher-education students, whereas our potential target group for EE-intervention would be young children. They would not perform at ceiling level at -5 dB SNR, and thus there could potentially be more EE-improvement. Alternatively, we may also consider the possibility that EE at lower noise levels might have a limited effect on outcome measures affecting speech intelligibility, but may influence processing speed or listening effort in participants with dyslexia considerably. Our experiment protocol does not allow us to provide evidence concerning this hypothesis. However, we do argue that future studies to specifically address long term effects of EE on both speech perception and moreover phonological abilities and literacy skills take into consideration alternative outcome measures as well.

Another important caveat is that typical readers and participants with dyslexia have different SRTs in the baseline condition. Thus, it could be possible that the interaction effect between group and EE-benefit merely reflects a ceiling effect in the typical reading group beyond a certain SNR. However, when comparing both speech conditions in our study, equal benefit from EE was found for typical readers, whereas the benefit was smaller for participants with dyslexia in the vocoded conditions. This suggests that the interaction is mainly driven by the performance of the participants with dyslexia, and not simply by the typical readers' possible ceiling effect. Nevertheless, it is not possible to truly reject the notion of a ceiling effect based on our results. However, we do not believe that this alters our overall conclusion since the main goal of our study was to assess whether enhancing onsets improves speech perception for participants with dyslexia to the level of typical readers' standard performance. Our results indeed suggest that around this

SNR (-10.2 dB), a ceiling effect does not yet appear to constrain the performance of the EE algorithm.

Other than the possibility of real-time implementation, a substantial benefit of EE is the feasibility to implement EE before initial reading instruction. Generally dyslexia is diagnosed relative late in childhood, after the most effective time for targeted intervention has already passed (Butterworth & Kovas 2013; Ozernov-Palchik & Gaab 2016). Once diagnosed, traditional remediation is instructional, and focuses on learning phonetic structures explicitly. Behavioral gains, together with increased activation in left temporo-parietal regions (typically involved during phonological processing), are found for both children and compensated adults after intensive instruction (Temple et al. 2003; Aylward et al. 2003; Eden et al. 2004). However, improvements are not typically found for all participants, are not necessarily maintained, and are not large enough for participants with dyslexia to completely catch up to typical readers (Gabrieli 2009). Our results show that EE benefits speech perception passively, and can be used before the outset of dyslexia. This allows preventive intervention during a time period where brain plasticity is most important, which could potentially strengthen the refinement of phoneme representations altogether.

The individual deviance profiles show that speech perception deficits might not be consistently associated with dyslexia and that consequently, EE intervention might not be effective for all individuals with dyslexia. However, this does not rule out the possibility that sensory impairments contribute to the reading deficits. First, we included a sample of high-functioning students and let them perform a task, which consisted of a situation encountered daily: speech understanding in noise. Therefore, we might expect a certain level of compensation for most of the dyslexic group. Still, we found clear speech perception deficits in half of the student with dyslexia (individual

deviance profiles: ~53%) in addition to impairments to multiple phonological abilities. Second, we found that speech perception deficits were not necessary for dyslexia, but neither were phonological difficulties (individual deviance profiles: ~85%), which have been assigned the 'core deficit' in dyslexia for the past decades (Goswami 2000; Snowling 2000). Considering that reading is a complex skill, and dyslexia is a rather heterogonous disorder, it is possible that dyslexia might manifest from different configurations of deficits including sensory impairments such as speech perception deficits (Rosner et al. 2003; Lallier et al. 2013).

Although the relationship between speech perception, phonological abilities and literacy has been proven to be highly complex (Joanisse et al. 2000; Robertson et al. 2009), we believe that our results highlight the necessity of auditory intervention. Future intervention studies are required to assess the direct and longitudinal impact of envelope enhancement on phonological abilities and literacy outcomes. Nevertheless, our findings provide a unique basis for early intervention and therapy by targeting the primary deficit in dyslexia without any requirement of explicit learning.

REFERENCES

Aylward, E., Richards, T., Berninger, V. (2003). Instructional treatment associated with changes in brain activation in children with dyslexia. *Neurology*, 61, 212–219.

Boets, B., Vandermosten, M., Poelmans, H., et al. (2011). Preschool impairments in auditory processing and speech perception uniquely predict future reading problems. *Res. Dev. Disabil.*, 32, 560–570.

Van den Bos, K.P., Spelberg, C.L.H., Scheepstra, J.M.A., et al. (1994). *De Klepel. Vorm A en B. Een test voor leesvaardigheid van pseudowoorden. Verantwoording, handleiding, diagnostiek en behandeling.*, Nijmegen: Pearson.

Brady, S., Shankweiler, D., Mann, V. (1983). Speech Perception and Memory Coding in Relation to Reading Ability. *J. Exp. Child Psychol.*, 35, 345–367.

Brus, B.T., Voeten, M.J.M. (1979). *Eén-Minuut-Test, vorm A en B. Verantwoording en Handleiding*, Nijmegen: Berkhout: Pearson.

Butterworth, B., Kovas, Y. (2013). Understanding neurocognitive developmental disorders can improve eduction for all. *Science*, 340, 300–305.

van Buuren, R.A. Van, Festen, J.M., Houtgast, T., et al. (1999). Compression and expansion of the temporal envelope : Evaluation of speech intelligibility and sound quality Compression and expansion of the temporal envelope : *J. Acoust. Soc. Am.*, 105, 2903–2913.

Calcus, A., Lorenzi, C., Collet, G., et al. (2016). Is there a relationship between speech indentification in noise and categorical perception in children with dyslexia. *Am. J. Speech-Language Pathol.*, 25, 1–15.

Clarkson, P.M., Bahgat, S.F., Clarkson, P.M., et al. (1991). Envelope expansion methods for speech enhancement Envelope expansion methods for speech enhancement. *J. Acoust. Soc. Am.*, 89, 1378–1382.

Doelling, K.B., Arnal, L.H., Ghitza, O., et al. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *Neuroimage*, 85, 761–768.

Dole, M., Hoen, M., Meunier, F. (2012). Speech-in-noise perception deficit in adults with

dyslexia: Effects of background type and listening configuration. *Neuropsychologia*, 50, 1543–1552.

Dorman, M.F., Loizou, P.C., Rainey, D. (1997). Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. *J. Acoust. Soc. Am.*, 102, 2403–2411.

Eden, G.F., Jones, K.M., Cappell, K., et al. (2004). Neural changes following remediation in adult developmental dyslexia. *Neuron*, 44, 411–422.

Fastl, H., Zwicker, E. (2007). *Psychoacoustics: Fast and Models* 3rd ed., Springer Science & Business Media.

Francart, T., van Wieringen, A., Wouters, J. (2008). APEX 3: a multi-purpose test platform for auditory psychophysical experiments. *J. Neurosci. Methods*, 172, 283–293.

Francart, T., Van Wieringen, A., Wouters, J. (2011). Comparison of fluctuating maskers for speech recognition tests. *Int. J. Audiol.*, 50, 2–13.

Friesen, L.M., Shannon, R. V, Baskent, D., et al. (2001). Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. *J. Acoust. Soc. Am.*, 110, 1150–1163.

Fu, Q., Shannon, R. V (1999). Recognition of spectrally degraded speech in noise with nonlinear amplitude mapping. *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process.*, 369–372.

Gabrieli, J.D.E. (2009). Dyslexia: A New Synergy Between Education and Dyslexia: A New Synergy Between Education and Cognitive Neuroscience. *Science*, 325, 280–284.

Ghesquière, P. (1998). *Algemene toets gevorderde spelling van het Nederlands (AT-GSN): Verantwoording en handleiding [unpublished]*, Leuven, Belgium: KU Leuven.

Goossens, T., Vercammen, C., Wouters, J., et al. (2017). Masked speech perception across the adult lifespan: Impact of age and hearing impairment. *Hear. Res.*, 344, 109–124.

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends Cogn. Sci.*, 15, 3–10.

Goswami, U. (2000). Phonological representations, reading development and dyslexia: Towards a cross-linguistic theoretical framework. *Dyslexia*, 6, 133–151.

Goswami, U. (2015). Sensory theories of developmental dyslexia: three challenges for research. *Nat Rev Neurosci*, 16, 43–54.

Goswami, U., Cumming, R., Chait, M., et al. (2016). Perception of filtered speech by children with developmental dyslexia and children with specific language impairments. *Front. Psychol.*, 7, 1–16.

Goswami, U., Gerson, D., Astruc, L. (2010). Amplitude envelope perception, phonology and prosodic sensitivity in children with developmental dyslexia. *Read. Writ.*, 23, 995–1019.

Goswami, U., Thomson, J., Richardson, U., et al. (2002). Amplitude envelope onsets and developmental dyslexia: A new hypothesis. *Proc. Natl. Acad. Sci.*, 99, 10911–10916.

Gross, J., Hoogenboom, N., Thut, G., et al. (2013). Speech Rhythms and Multiplexed Oscillatory Sensory Coding in the Human Brain. *PLOS Biol.*, 11, 1–14.

Hämäläinen, J.A., Rupp, A., Soltész, F., et al. (2012). Reduced phase locking to slow amplitude modulation in adults with dyslexia: An MEG study. *Neuroimage*, 59, 2952–2961.

Hazan, V., Messaoud-Galusi, S., Rosen, S., et al. (2009). Speech Perception Abilities of Adults With Dyslexia: Is There Any Evidence for a True Deficit? *J. Speech, Lang. Hear. Res.*, 52, 1510–1529.

Hornickel, J., Zecker, S.G., Bradlow, A.R., et al. (2012). Assistive listening devices drive neuroplasticity in children with dyslexia. *Proc. Natl. Acad. Sci.*, 109, 16731–16736.

Joanisse, M.F., Manis, F.R., Keating, P., et al. (2000). Language deficits in dyslexic children: Speech perception, phonology, and morphology. *J. Exp. Child Psychol.*, 77, 30–60.

Johnson, E.P., Pennington, B.F., Lowenstein, J.H., et al. (2011). Sensitivity to structure in the speech signal by children with speech sound disorder and reading disability. *J. Commun. Disord.*, 44, 294–314.

Koning, R., Wouters, J. (2012). The potential of onset enhancement for increased speech intelligibility in auditory prostheses. *J. Acoust. Soc. Am.*, 132, 2569–2581.

Lallier, M., Thierry, G., Tainturier, M.J. (2013). On the importance of considering individual profiles when investigating the role of auditory sequential deficits in developmental dyslexia. *Cognition*, 126, 121–127.

Langhans, T., Strube, H.W. (1982). Speech enhancement by nonlinear multiband envelope filtering. *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process.*, 156–159.

Leong, V., Hämäläinen, J., Soltész, F., et al. (2011). Rise time perception and detection of syllable stress in adults with developmental dyslexia. *J. Mem. Lang.*, 64, 59–73.

Lorenzi, C., Dumont, A., Füllgrabe, C. (2000). Use of Temporal Envelope Cues by Children With Developmental Dyslexia. *J. Speech Lang. Hear. Res.*, 43, 1367–1379.

Menell, P., McAnally, K.I., Stein, J.F. (1999). Psychophysical Sensitivity and Physiological Response to Amplitude Modulation in Adult Dyslexic Listeners. *J. Speech Lang. Hear. Res.*, 42, 797–803.

Messaoud-Galusi, S., Hazan, V., Rosen, S. (2011). Investigating Speech Perception in Children With Dyslexia: Is There Evidence of a Consistent Deficit in Individuals? *J. Speech Lang. Hear. Res.*, 54, 1682–1701.

Molinaro, N., Lizarazu, M., Lallier, M., et al. (2016). Out-of-synchrony speech entrainment in developmental dyslexia. *Hum. Brain Mapp.*, 37, 2767–2783.

Nittrouer, S., Krieg, L.M., Lowenstein, J.H. (2018). Speech Recognition in Noise by Children with and without Dyslexia: How is it Related to Reading? *Res. Dev. Disabil.*, 77, 98–113.

Nittrouer, S., Lowenstein, J.H. (2013). Perceptual organization of speech signals by children with and without dyslexia. *Res. Dev. Disabil.*, 34, 2304–2325.

Ozernov-Palchik, O., Gaab, N. (2016). Tackling the "dyslexia paradox": Reading brain and behavior for early markers of developmental dyslexiax. *Wiley Interdiscip. Rev. Cogn. Sci.*, 7, 156–176.

Pearsons, K., Bennett, R., Fidell, S. (1977). *Speech levels in various noise environments*, Washington, D.C.

Plomp, R. (1988). The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function. *J. Acoust. Soc. Am.*, 83, 2322–2327.

Poelmans, H., Luts, H., Vandermosten, M., et al. (2011). Reduced sensitivity to slow-rate dynamic auditory information in children with dyslexia. *Res. Dev. Disabil.*, 32, 2810–2819.

Power, A.J., Colling, L.J., Mead, N., et al. (2016). Neural encoding of the speech envelope by

children with developmental dyslexia. *Brain Lang.*, 160, 1–10.

R Core Team (2015). R: A language and environment for statistical computing.

Ramus, F. (2001). Talk of two theories. *Nature*, 412, 393–395.

Ramus, F., Rosen, S., Dakin, S.C., et al. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain*, 126, 841–865.

Richardson, U., Thomson, J.M., Scott, S.K., et al. (2004). Auditory processing skills and phonological representation in Dyslexic children. *Dyslexia*, 10, 215–233.

Robertson, E.K., Joanisse, M.F., Desroches, A.S., et al. (2009). Categorical speech perception deficits distinguish language and reading impairments in children. *Dev. Sci.*, 12, 753–767.

Rosen, S., Zhang, Y., Speers, K. (2015). Spectral density affects the intelligibility of tone-vocoded speech: Implications for cochlear implant simulations. *J. Acoust. Soc. Am.*, 138, EL318-EL323.

Rosner, B.S., Talcott, J.B., Hogg, J.D., et al. (2003). The Perception of "Sine-Wave Speech " by Adults With Developmental Dyslexia. *J. Speech Lang. Hear. Res.*, 46, 68–79.

Schroeder, M.R. (1981). Modulation transfer functions: definition and measurement. *Acta Acust. united with Acust.*, 49, 179–182.

Shannon, R. V, Fu, Q.J., Galvin, J. (2004). The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. *Acta Otolaryngol.*, 124, 50–54.

Shannon, R. V, Zeng, F.G., Kamath, V., et al. (1995). Speech recognition with primarily temporal cues. *Science (80-.).*, 270, 303–304.

Smeds, K., Wolters, F., Rung, M. (2015). Estimation of Signal-to-Noise Ratios in Realistic Sound Scenarios. *J. Am. Acad. Audiol.*, 26, 183–196.

Smits, C., Kapteyn, T.S., Houtgast, T. (2004). Development and validation of an automatic speech-in-noise screening test by telephone. *Int. J. Audiol.*, 43, 15–28.

Snowling, M.J. (2000). *Dyslexia* 2nd ed., Oxford, UK: Blackwell.

Soltész, F., Szucs, D., Leong, V., et al. (2013). Differential Entrainment of Neuroelectric Delta

Oscillations in Developmental Dyslexia. *PLoS One*, 8, e76608.

Sperling, A.J., Lu, Z.L., Manis, F.R., et al. (2005). Deficits in perceptual noise exclusion in developmental dyslexia. *Nat. Neurosci.*, 8, 862–863.

Stilp, C.E., Goupell, M.J. (2015). Spectral and temporal resolutions of information-bearing acoustic changes for understanding vocoded sentences. *J. Acoust. Soc. Am.*, 137, 844–855.

Temple, E., Deutsch, G.K., Poldrack, R.A., et al. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: Evidence from functional MRI. *Proc. Natl. Acad. Sci.*, 100, 2860–2865.

The MathWorks Inc (2013). MATLAB and Statistics Toolbox Release 2013a.

Vanvooren, S., Poelmans, H., De Vos, A., et al. (2017). Do prereaders' auditory processing and speech perception predict later literacy? *Res. Dev. Disabil.*, 70, 138–151.

Varnet, L., Ortiz-Barajas, M.C., Erra, R.G., et al. (2017). A cross-linguistic study of speech modulation spectra. *J. Acoust. Soc. Am.*, 142, 1976–1989.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale* 3rd ed., San Antonio, Texas: The Psychological Corporation.

Van Wieringen, A., Wouters, J. (2008). LIST and LINT: Sentences and numbers for quantifying speech understanding in severely impaired listeners for Flanders and the Netherlands. *Int. J. Audiol.*, 47, 348–355.

Witton, C., Stein, J.F., Stoodley, C.J., et al. (2002). Separate Influences of Acoustic AM and FM Sensitivity on the Phonological Decoding Skills of Impaired and Normal Readers. *J. Cogn. Neurosci.*, 14, 866–874.

Ziegler, J.C., Pech-Georgel, C., George, F., et al. (2009). Speech-perception-in-noise deficits in dyslexia. *Dev. Sci.*, 12, 732–745.

Acknowledgments

We are grateful to all our participants who made time to participate in this study. Our special thanks goes to Robin Gransier for his assistance and insightful comments with regard to data analysis.

This work was funded by the Research Foundation - Flanders (FWO) (G.0A91.15N).

Portions of this article were presented at the The International Hearing Aid Research Conference, Lake Tahoe, California (15-18 August 2018), and at the Audiological Research Centers in Europe (ARCHES) Conference, Nottingham, UK (12-13 November 2018).

Regarding author contributions, T.V.H. and J.W. designed the experiments. T.V.H. and P.G. designed the cognitive tasks and recruited the subjects. T.V.H. performed the research, analyzed the data and wrote the manuscript. A.M.T. and J.W. worked out and optimized the signal processing. P.G. and J.W. supervised the research. All authors edited the manuscript.

Address for correspondence: Tilde Van Hirtum, O&N2, Herestraat 49 – Box 721, B-3000 Leuven. E-mail: tilde.vanhirtum@kuleuven.be

Figure legends

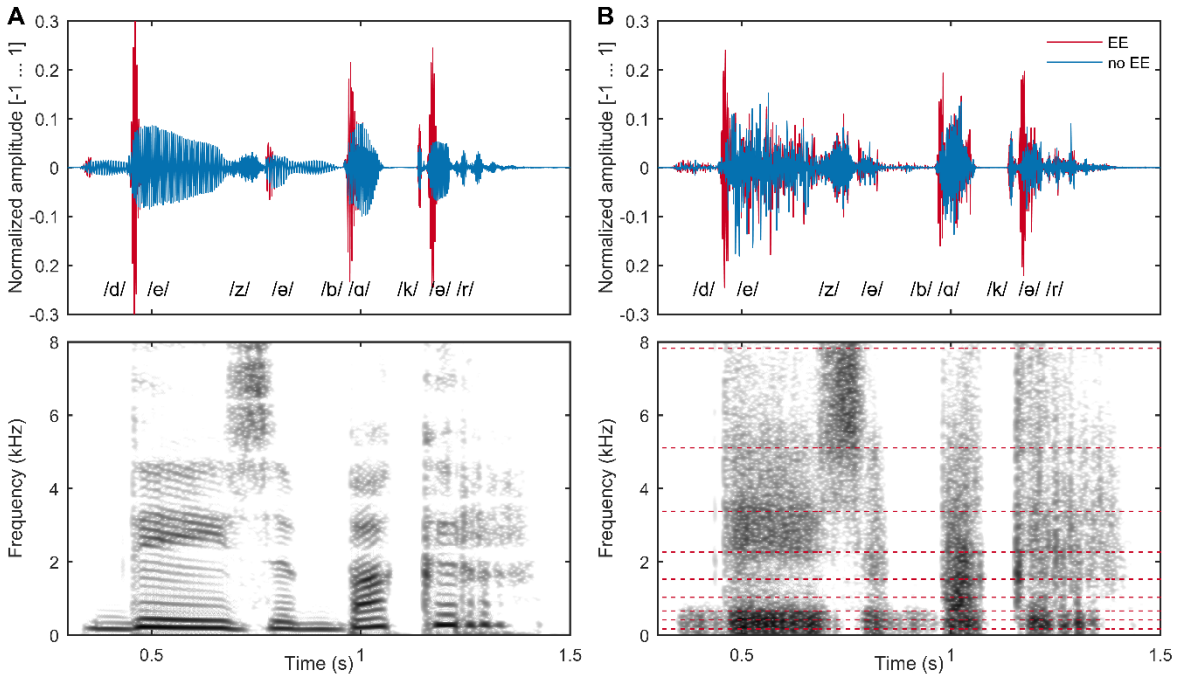

Figure 1. Stimuli. A, Waveform (top) of part of one natural speech sentence without envelope enhancement (EE, blue) and with EE (red), and spectrogram (bottom) of the same example without EE. **B,** Waveform (top) of part of one noise-band vocoded sentence (featuring degraded spectral resolution and limited temporal fine structure) without EE (blue) and with EE (red), and spectrogram (bottom) of the same example without EE. Dashed lines in B (bottom panel) represent the cut-off frequencies of the analysis filter.

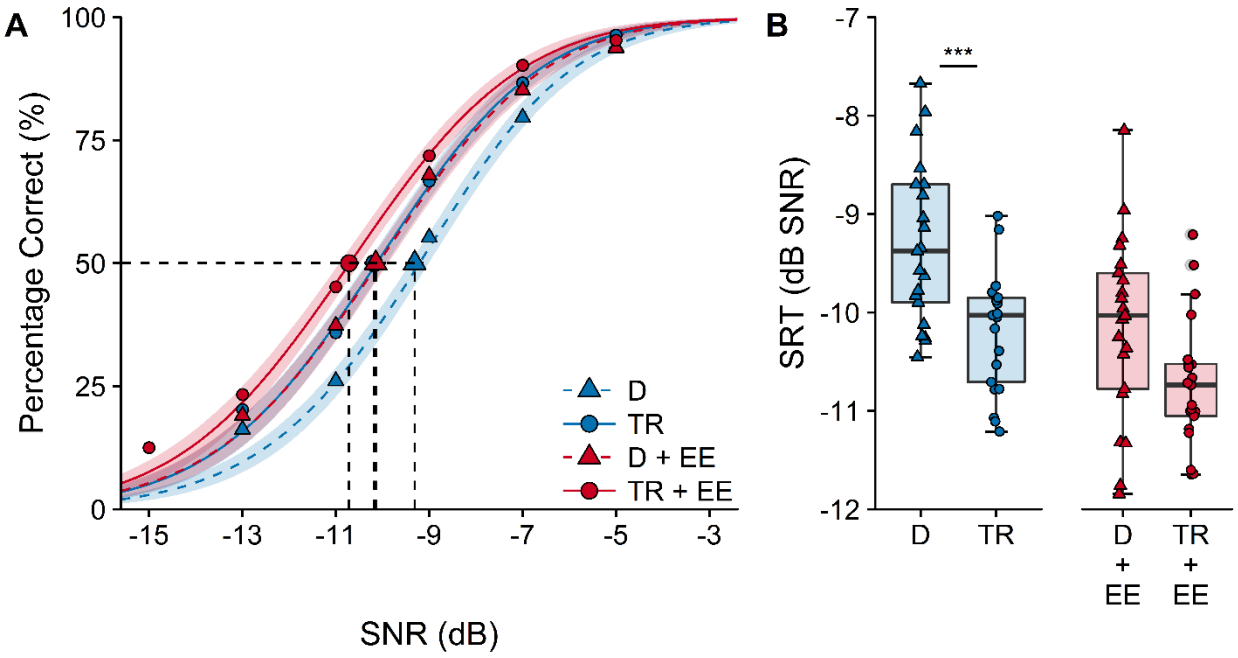

Figure 2. Speech reception thresholds for natural speech. A, Psychometric functions fitted using a cumulative Gaussian distribution, showing natural speech without envelope enhancement (EE; blue) and with EE (red). Triangles and circles represent the average raw scores for participants with dyslexia (D) and typical readers (TR) respectively. Dashed lines represent the average SRTs (speech reception thresholds) expressed as speech-in-noise ratios (SNR; in dB). The transparent ribbon represents the psychometric function ± the error of fit of the model. **B,** SRTs for natural speech without EE (left, blue) and with EE (right, red) for participants with dyslexia (D) and typical readers (TR). Individual data points in each boxplot represent individual participants, to show variation. Lines indicate the median value, and whiskers are 1.5 x IQR. *** $p < 0.001$.

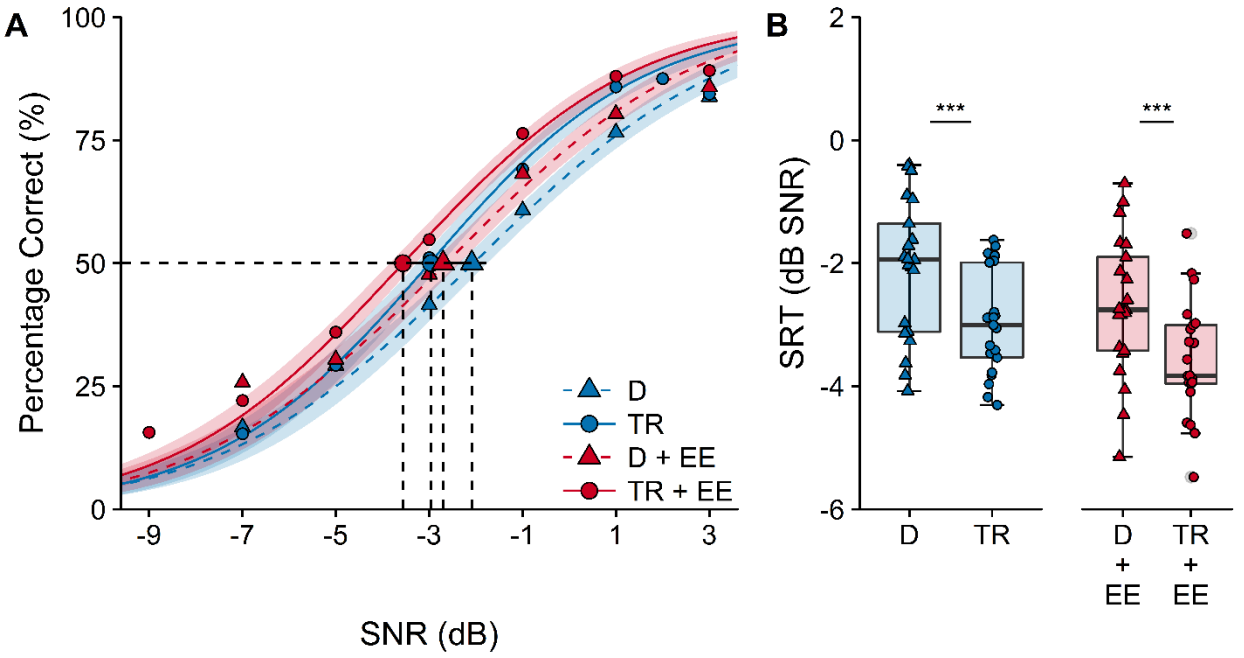

Figure 3. Speech reception thresholds for noise-band vocoded speech. A, Psychometric functions fitted using a cumulative Gaussian distribution, showing vocoded speech without envelope enhancement (EE, blue) and with EE (red). Triangles and circles represent the average raw scores for participants with dyslexia (D) and typical readers (TR) respectively. Dashed lines represent the average SRTs (speech reception thresholds) expressed as speech-in-noise ratios (SNR; in dB). The transparent ribbon represents the psychometric function ± the error of fit of the model. **B,** SRTs for vocoded speech without EE (left, blue) and with EE (right, red) for participants with dyslexia (D) and typical readers (TR). Individual data points in each boxplot represent individual participants, to show variation. Lines indicate the median value, and whiskers are 1.5 x IQR. *** $p < 0.001$.

Tables

Table 1: Performance of typical readers (TR) and participants with dyslexia (D) in the various literacy and cognitive tasks. Standard deviations are shown in parentheses.

	TR	D		
	M (SD)	M (SD)	*P*	$r^{\dagger}$
Literacy				
EMT (# correct/min)	108.8 (6.9)	71.9 (10.8)	< 0.001	0.91
Klepel (# correct/2min)	105.9 (9.0)	66.3 (13.6)	< 0.001	0.88
Spelling (# errors)	10.1 (4.8)	28.5 (8.1)	< 0.001	0.84
Composite z-score	*0 (0.9)*	*-4.5 (1.2)*	*< 0.001*	*0.92*
Intelligence				
Non-verbal*	111.3 (9.4)	113.3 (13.2)	0.578	0.09
Language				
Vocabulary*	115.9 (7.5)	109.8 (6.2)	0.006	0.42
Phonological Awareness				
Accuracy	0.88 (0.1)	0.68 (0.1)	< 0.001	0.67
Response time (s)	3.6 (1.3)	7.4 (3.7)	< 0.001	0.66
Composite z-score	*0 (0.4)*	*-2.5 (1.5)*	*< 0.001*	*0.73*
Phonological Fluency				
RAN objects (s)	28.2 (2.9)	35.3 (4.7)	< 0.001	0.72
RAN colours (s)	24.5 (2.5)	30.2 (4.7)	< 0.001	0.67
RAN digits (s)	15.8 (1.8)	19.7 (3.9)	< 0.001	0.62
RAN letters (s)	15.7 (2.1)	23 (7.26)	< 0.001	0.67
Composite z-score	*0 (0.4)*	*-2.6 (1.0)*	*< 0.001*	*0.77*
Verbal short-term memory				
Digit span forward (# correct)	10.4 (1.8)	8.2 (1.6)	< 0.001	0.55
Working memory				
Digit span backward (# correct)	7.9 (2.1)	7.1 (2.1)	0.24	0.18

Note: Mean (M), standard deviations (SD), *p* values and effect sizes (*r*) of group differences from independent t-test are provided for each subtest. For literacy phonological awareness and phonological fluency tasks, raw scores were converted to z-scores and averaged to obtain a composite z-score. The raw scores of spelling, response time (phonological awareness) and RAN tasks (phonological fluency) were first multiplied by -1 so that all z-scores were interpretable in the same direction.

*Standardized IQ test (M = 100, SD = 15)

†Effect sizes (Pearson's *r*) were calculated using the following formula: $r = \sqrt{t^2/(t^2 + df)}$

Table 2: Speech perception performance depicted as average SRTs (dB), their precision values and standard deviations (SD) for typical readers (TR) and participants with dyslexia (D)

	TR			D		
	SRT (dB)	Precision (dB)	SD (dB)	SRT (dB)	Precision (dB)	SD (dB)
Unprocessed speech						
Without EE	-10.19	0.22	0.60	-9.31	0.20	0.81
EE	-10.74	0.22	0.65	-10.15	0.22	0.93
Vocoded speech						
Without EE	-2.97	0.30	0.85	-2.09	0.31	1.13
EE	-3.57	0.30	0.94	-2.70	0.31	1.14

Table 3: Stepwise regression analyses of literacy performance for unprocessed speech. Cognitive control variables are entered in step 1, whereas speech perception thresholds (SRTs) for unprocessed speech are entered in step 2.

		Literacy		
Controlling for	Variables	R^2	R^2 change	F change
Intelligence	Step 1 : non-verbal IQ	-0.02		0.49
	Step2: SRT	0.28	0.3	16.31***
Language	Step 1: Vocabulary	0.2		13.66***
	Step 2: SRT	0.34	0.14	9.81**
Phonological awareness	Step 1: Spoonerisms	0.43		34.44***
	Step 2: SRT	0.47	0.04	4.19 *
Phonological fluency	Step 1: RAN	0.45		38.08***
	Step 2: SRT	0.50	0.05	5.22*
Verbal short-term memory	Step 1: Digit span forward	0.29		19.53***
	Step 2: SRT	0.36	0.07	5.58*
Working memory	Step 1: Digit span backward	0.01		ns
	Step 2: SRT	0.27	0.26	15.37***

Note: R^2 = adjusted R, *** $p < 0.001$, ** $p < 0.01$, *$p < 0.05$, . $p < 0.1$, ns = not significant

Table 4: Stepwise regression analyses of literacy performance for vocoded speech. Cognitive control variables are entered in step 1, whereas speech perception thresholds (SRTs) for vocoded speech are entered in step 2.

		Literacy		
Controlling for	Variables	R^2	R^2 change	F change
Intelligence	Step 1 : non-verbal IQ	-0.02		0.40
	Step2: SRT	0.11	0.13	6.46 *
Language	Step 1: Vocabulary	0.2		12.83***
	Step 2: SRT	0.30	0.10	6.86*
Phonological awareness	Step 1: Spoonerisms	0.43		33.5 ***
	Step 2: SRT	0.46	0.03	3.04 .
Phonological fluency	Step 1: RAN	0.45		34.87***
	Step 2: SRT	0.46	0.01	1.49
Verbal short-term memory	Step 1: Digit span forward	0.29		18.09***
	Step 2: SRT	0.31	0.02	ns
Working memory	Step 1: Digit span backward	0.01		ns
	Step 2: SRT	0.15	0.15	7.86 **

Note: R^2 = adjusted R, *** $p < 0.001$, ** $p < 0.01$, *$p < 0.05$, . $p < 0.1$, ns = not significant