.

Informational masking of speech in dyslexic children

Axelle CalcusCécile ColinPaul DeltenreRégine KolinskyQJF

Articles you may be interested in

Voice source characterization using pitch synchronous discrete cosine transform for speaker identification

The Journal of the Acoustical Society of America **137**, EL469 (2015); 10.1121/1.4921679

Informational masking of speech in dyslexic children

Axelle Calcus
Fonds de la Recherche Scientifique—FNRS (FRS-FNRS), Unité de Recherche en Neurosciences Cognitives (UNESCOG), Center for Research in Cognition and Neurosciences (CRCN), Université Libre de Bruxelles (ULB), Brussels, Belgium
acalcus@ulb.ac.be

Cécile Colin
Unité de Recherche en Neurosciences Cognitives (UNESCOG), Center for Research in Cognition and Neurosciences (CRCN), Université Libre de Bruxelles, Brussels, Belgium
ccolin@ulb.ac.be

Paul Deltenre
Laboratoire de Neurophysiologie Sensorielle et Cognitive, Hôpital Brugmann, Brussels, Belgium
paul.deltenre@chu-brugmann.be

Régine Kolinsky
Fonds de la Recherche Scientifique—FNRS (FRS-FNRS), Unité de Recherche en Neurosciences Cognitives (UNESCOG), Center for Research in Cognition and Neurosciences (CRCN), Université Libre de Bruxelles (ULB), Brussels, Belgium
rkolins@ulb.ac.be

Abstract: Studies evaluating speech perception in noise have reported inconsistent results regarding a potential deficit in dyslexic children. So far, most of them investigated energetic masking. The present study evaluated situations inducing mostly informational masking, which reflects cognitive interference induced by the masker. Dyslexic children were asked to identify a female target syllable presented in quiet, babble, unmodulated, and modulated speech-shaped noise. Whereas their performance was comparable to normal-reading children in quiet, it dropped significantly in all noisy conditions compared to age-, but not reading level-matched controls. Interestingly, noise affected similarly the reception of voicing, place, and manner of articulation in dyslexic and normal-reading children.
© 2015 Acoustical Society of America
[QJF]

1. Introduction

Dyslexia is a neurodevelopmental disorder affecting reading acquisition in about 7% of school age children, despite adequate intelligence, sensory abilities, and educational opportunities (Snowling, 2000). According to the phonological hypothesis, poor phonological representations and/or access to phonological information is likely to impede reading acquisition (Vellutino *et al.*, 2004). Consistently, speech perception deficits have often been reported in dyslexics (e.g., Bogliotti *et al.*, 2008) and have been suggested to induce poor phonological abilities, hence impeding acquisition of the phoneme-grapheme conversion, ultimately leading to reading disorders. However, inconsistencies in the observed speech perception deficits have led some researchers to question the impact of the listening conditions.

Whereas optimal listening environments provide listeners with highly redundant acoustic information about the speech signal (Zeng *et al.*, 2005), the presence of noise in suboptimal backgrounds impedes speech perception by degrading representation and/or

limiting access to some acoustical cues. Therefore, weak and/or underspecified representations of speech sounds that might remain unnoticed in optimal listening environments could reveal themselves in noisy backgrounds. This idea has been explored in dyslexics, using various types of material. In a consonant identification task, Ziegler *et al.* (2009) showed that dyslexic children's performance was comparable to controls' in quiet, but poorer in stationary speech-shaped noise (SSN). However, they benefited from the "valleys" of a fluctuating background noise to experience normal masking release, an ability that requires good spectro-temporal resolution at the peripheral level. Interestingly, dyslexics' deficit was even observed when compared to younger, reading age-matched control children, suggesting that their speech perception deficit was not related to a mere delay in reading acquisition, but constituted a core difficulty inherent to dyslexia. Later studies replicated this finding using either vowel (Poelmans *et al.*, 2011) or sentence (Chandrasekaran *et al.*, 2009) identification tasks presented in a SSN background. However, Dole *et al.* (2012) showed that dyslexic adults exhibited preserved perception of speech under SSN, but were impaired when presented with babble noise. Later studies revealed a small but significant deficit of speech discrimination (but not identification) in babble noise (Messaoud-Galusi *et al.*, 2011; Hazan *et al.*, 2013).

Different auditory backgrounds are likely to induce different types of masking interference. Energetic masking (EM) arises because of a spectral interference between target and masker falling within the same auditory filter. It has been widely associated with the masking effects of a "steady" noise (e.g., SSN), despite recent evidence highlighting the important contribution of random amplitude modulations in the SSN (i.e., modulation masking; see Stone *et al.*, 2012). Informational masking (IM) has been equated to non-EM (Durlach *et al.*, 2003), and was initially observed in the absence of any spectral overlap between target and masker pure tones (Neff and Green, 1987). Later studies revealed that EM could not account for all the difficulty induced by a background of simultaneous talkers (i.e., babble noise, see Brungart *et al.*, 2001). Therefore, in the context of speech, IM is defined as the excess of masking that cannot be explained by the spectral interference between target and masker. Stimulus uncertainty and/or similarity have been demonstrated to influence IM (Durlach *et al.*, 2003), as they reflect a failure of object-based selection, hence preventing the listeners to perform the auditory scene analysis (Shinn-Cunningham, 2008). Overall, whereas EM arises because of frequency selectivity limitations at the peripheral level, IM reflects processing capacity limitations at a more central level. Therefore, it is crucial to understand the specific contribution of peripheral and central interference to the difficulty encountered by dyslexic children perceiving speech in noise, as they would reflect a failure at very different levels of processing of the speech signal.

The nature of the background noise is thus likely to influence the outcome of studies investigating speech intelligibility in dyslexics. Very few studies specifically investigated IM in this population. Our recent work suggests that dyslexic children experience difficulties in complex tone sequences inducing pure IM in comparison to both reading level- and age-matched controls (Calcus *et al.*, 2015). In contrast, Messaoud-Galusi *et al.* (2011) and Hazan *et al.* (2013) failed to evidence a deficit in dyslexics presented with a babble noise background. Yet, as there was no attempt to remove its energetic component, the babble noise used in those two studies simultaneously induced EM and IM.

Therefore, the aim of the present experiment was to investigate dyslexics' sensitivity to noisy situations minimizing EM but inducing various amounts of IM. To do so, we compared intelligibility of speech presented in three different backgrounds: babble, envelope-modulated speech-shaped noise (henceforth, *eSSN*), and unmodulated SSN. Whereas the babble was expected to induce an important cognitive interference, eSSN and SSN aimed at investigating the impact of spectral interference respectively with or without low-amplitude valleys. In order to minimize cochlear EM, target and maskers were presented dichotically. Because dyslexics were previously shown to exhibit a failure of unilateral selective attention with nonspeech sounds (Smith and Griffiths, 1987; Calcus *et al.*, 2015), the degree of predictability of the target lateralization was varied.

We hypothesized that a predictable target lateralization would provide children with helpful spatial lateralization cues that would help them selectively focus on the relevant target. However, a failure in selective auditory attention might already impede target identification in the predictable condition in dyslexic children.

In addition, we performed information transmission analyses (Miller and Niceley, 1955) on the basis of individual confusion matrices obtained across all noise conditions in order to evaluate the specific reception of voicing, place of articulation, and manner. Indeed, phonetic feature analyses have led to contradictory results, some suggesting that reception of place was specifically impaired in dyslexic children (Ziegler *et al.*, 2009), while others pointed at reception of voicing (Hazan *et al.*, 2013). Altogether, determining if the reception of a phonetic trait is specifically affected in noisy backgrounds might be of crucial importance in refining clinical treatment of dyslexic children.

2. Method

2.1 Participants

Children with phonological dyslexia ($n = 10$) were included in the study if they were free of other developmental disorder (e.g., oral speech impairment, attention deficit and hyperactivity disorder, autism), and if their reading level was at least 2 SD (standard deviation) below the norm on MIM and REGUL reading tests (BELEC, Mousty and Leybaert, 1999) that evaluate reading of regular, irregular, and pseudo-words varying in complexity. A pseudo-word repetition task and a phoneme deletion task evaluated phonological processing. Two groups of control children were included in the study. They were matched to the dyslexics either on chronological age ($n = 10$) or reading level ($n = 10$). Table 1 summarizes the results of the ancillary tests for each group. All children had normal audiometric thresholds as measured at octave intervals from 0.25 to 8 kHz. Their performance IQ was above 80 on the Weschler Non Verbal scale. The study was conducted with the understanding and consent of all children and their parents.

2.2 Material

A set of 64 natural Consonant-Vowel (CV) stimuli was recorded, V being always /a/, and C being chosen among the /p,t,k,b,d,g,f,s,ʃ,m,n,r,l,v,z,j/ consonant set. Two French native female speakers each recorded 2 exemplars of the 16 possible syllables in a soundproof booth (mean duration: 286 ms; SD: 66). Signal was digitized via a 16-bit A/D converter at 44.1 kHz sampling frequency. CV identification was assessed in quiet and in three different noisy backgrounds: natural babble, eSSN, and unmodulated SSN. Babble stimuli consisted of a mixture of eight male speakers. Each talker was first recorded in a sound-proof booth while reading extracts of French press. Individual recordings were edited in order to remove silences longer than 1 s, pronunciation errors, and proper nouns, resulting in sound files approximately 90 s in duration

Table 1. Characteristics of the participants (dyslexics: DYS; reading-level matched controls: RL; age-matched controls: AGE), and mean scores in the ancillary tests (standard deviations in parentheses). The last four columns present the results of independent-samples *t*-tests (with group as a between-subjects variable).

	Groups			DYS vs RL		DYS vs AGE	
	DYS (n = 10)	RL (n = 10)	AGE (n = 10)	*t*	*p*	*t*	*p*
Chronological age	11.2 (0.85)	8.5 (1.1)	11.2 (0.76)	5.7	0.001	−0.29	0.76
Sex (male)	8	3	3				
Non Verbal IQ	101.2 (13.9)	103.4 (7.4)	102.3 (13.7)	−0.44	0.66	−0.18	0.86
Reading							
MIM	57.7 (6.1)	61.6 (5.0)	64.7 (5.0)	−1.55	0.14	−2.79	0.01
REGUL	37.2 (3.4)	38.5 (4.6)	45.7 (2.8)	−0.707	0.48	−6.01	0.001
Metaphonology	58.4 (5.9)	65.0 (4.1)	66.6 (5.4)	−2.92	0.009	−3.24	0.004

for each talker, that were normalized to a common root-mean-square amplitude. Both eSSN and SSN were derived from the babble noise using custom-made MATLAB (MathWorks, Natick, MA) programs. The spectrum of the original signal was computed using a fast Fourier transformation. A new signal having the equivalent power spectrum but randomized phases was generated in order to create the SSN. The eSSN signal was then constructed by multiplying the envelope of the original babble (obtained by 60 Hz low-pass filtering of the full-wave rectified signal) against the SSN. For every CV item, noise extracts were randomly selected from the 90 s initial waveform. Duration of the noise was adapted to match exactly that of the target syllable.

Target and maskers were presented dichotically, with target lateralization being either predictable or unpredictable. In the unpredictable condition, half of the targets were presented to the right ear, the others to the left ear, and their lateralization was randomly determined from trial to trial. Similarly, in the quiet condition, target syllables were presented monotically, and their lateralization changed randomly from trial to trial. In the predictable lateralization condition, target syllables always occurred in the right ear. Overall levels were calibrated to produce an average output level of 70 dB(A) for each background noise. Target syllables were presented at 40 dB(A). Piloting the experiment on three children showed that this extreme signal-to-noise ratio of -30 dB would avoid ceiling effects due to spatial lateralization inherent to dichotic presentation.

2.3 Procedures

Children were tested individually in a quiet room, over 2 sessions of about 1 h each. During the speech in noise task, the children were asked to focus on the female voice, and identify the stimulus by repeating what they heard. The experimenter encoded the responses; no feedback was given. Children were presented with a total of seven listening conditions. They started with the quiet condition, followed by the three noisy backgrounds presented in the unpredictable target lateralization condition, then with the three noisy backgrounds presented in the predictable condition. They were explicitly told that in the latter cases, target would always occur in the right ear, which they were encouraged to pay attention to. Presentation order of each noisy background (babble, eSSN, and SSN) was randomized within both lateralization predictability conditions.

3. Results

3.1 Speech identification in noise

Children's performance was evaluated in terms of percentage of correct identification in each listening condition (see Fig. 1). Given that the data in the quiet condition are binomial with scores near ceiling, a mixed levels regression of performance on the group was computed. The listeners' group was not a significant predictor of correct response [$\chi^2(14) = 16.12$, $p > 0.10$].

A repeated-measures analysis of variance (ANOVA) was then performed on percent correct identification in noise, with lateralization predictability (predictable vs unpredictable) and background noise (babble, eSSN, SSN) as within-subject factors, and group (dyslexic, reading level- and age matched-controls) as a between-subject factor. Bonferroni-corrected *post hoc t*-tests were used to specify main effects when necessary. Unsurprisingly, there was a significant main effect of lateralization predictability [$F(1,27) = 19.7$, $p < 0.001$, $\eta^2 = 0.05$], with better performance in the predictable compared to the unpredictable condition. There was also a significant main effect of the nature of the background noise [$F(2,54) = 42.9$, $p < 0.001$, $\eta^2 = 0.11$], without interaction with lateralization predictability [$F(2,54) = 1.64$, $p > 0.10$, $\eta^2 = 0.004$]. Overall, there was no significant difference between the SSN and eSSN conditions ($p > 0.50$), which were both significantly better performed than the babble condition (both $ps < 0.001$). Importantly, there was also a significant main effect of group [$F(2,27) = 5.58$, $p < 0.01$, $\eta^2 = 0.18$]: dyslexic children performed significantly worse than age-matched controls ($p < 0.05$), but did not differ from reading level-matched children ($p > 0.50$). There was
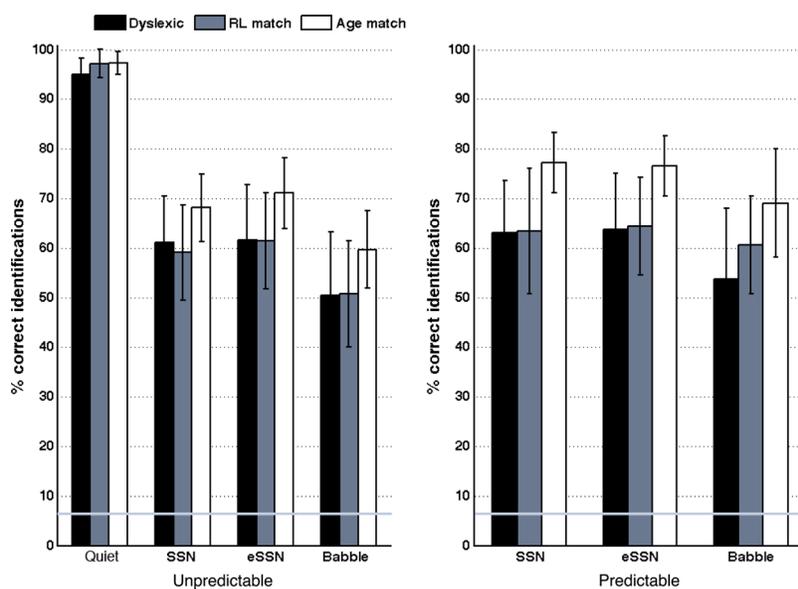
Fig. 1. (Color online) Percentage correct identification of a CV target in both unpredictable (left panel: quiet, SSN, eSSN, and babble) and predictable (SSN, eSSN, and babble) lateralization. Error bars represent standard deviation. The horizontal lines at the bottom represent chance performance.

no significant group $\times$ background noise or group $\times$ lateralization predictability interaction (both $ps > 0.10$, $\eta^2 < 0.01$). The triple interaction was not significant either ($p > 0.10$, $\eta^2 = 0.002$).

Last, we compared each group's performance in quiet versus noise. However, because performance in quiet was only investigated in the unpredictable target lateralization, we performed weighted planned contrasts (quiet vs three noise conditions). The results revealed a significant group $\times$ listening condition interaction [$F(1,27) = 5.22$, $p < 0.05$]: compared to both dyslexics and reading-level-matched controls, age-matched controls performed similarly in quiet (both $ps > 0.10$), but better in all three noise conditions (all $ps < 0.05$).

3.2 Phonetic feature transmission

We computed an ANOVA on the specific reception scores of three phonetic features, with the same factors as in the former ANOVA (group; lateralization predictability; background noise), plus phonetic feature (voicing, place, manner) as a within-subject factor. The results confirmed the significant main effects of noise, listening configuration, and group that were observed on overall performance [$F(2,54) = 36.6$, $p < 0.001$, $\eta^2 = 0.028$ and $F(1,27) = 25.2$, $p < 0.001$, $\eta^2 = 0.018$, $F(2,27) = 5.45$, $p < 0.05$, $\eta^2 = 0.03$, respectively]. Furthermore, we observed a significant effect of phonetic feature [$F(2,54) = 464.5$, $p < 0.001$, $\eta^2 = 0.63$]: place ($M = 80.8$, SD $= 17.05$) was better transmitted than both voicing ($M = 34.8$, SD $= 16.7$) and manner ($M = 36.05$, SD $= 14.9$; both $ps < 0.001$), which did not significantly differ from each other ($p > 0.50$). Crucially, phonetic feature did not interact with any other factor (all $ps > 0.05$, $\eta^2 < 0.01$).

3.3 Correlations with literacy-related performance

A potential link between auditory processing and language skills was evaluated by performing correlations between performance in each listening condition and both reading and metaphonological abilities in dyslexic children. There was no significant partial correlation between these variables when using age as a controlling variable (all $ps > 0.10$).

4. Discussion

Consistent with several studies (Chandrasekaran *et al.*, 2009; Ziegler *et al.*, 2009; Poelmans *et al.*, 2011), dyslexics' performance was comparable to controls' when the stimuli were presented in quiet, but was significantly worse than that of age-matched controls in all noisy backgrounds. This deficit held true for the three consonant features examined here, which tempers previous finding of a specific impairment of reception of voicing and/or place of articulation in noise in dyslexic individuals (Hazan *et al.*, 2013; Ziegler *et al.*, 2009). Dichotic presentation of the speech target and masker minimized cochlear EM while preserving IM. Listeners' sensitivity to IM has been shown to be mostly influenced by cognitive factors (Durlach *et al.*, 2003). Therefore, our main result underscores the importance of a central, cognitive contribution to the difficulty encountered by dyslexic children in ecological auditory backgrounds.

The various noisy backgrounds affected listeners' perception differently. There was no significant difference in performance under SSN compared to eSSN noise background. Fairly short tokens containing very limited envelope fluctuations likely limited the possibility of "central" masking release that would happen if listeners extracted information regarding the speech signal in the valleys after combining information from both ears. Yet, performance was significantly better in both SSN and eSSN than in the eight-talker babble background. This finding is consistent with a previous observation of a significant drop in adult listeners' consonant identification in an eight-talker babble compared to an eSSN background noise (Simpson and Cooke, 2005). Using diotic presentation, Brungart *et al.* (2001) showed that babble noise induced an excess of masking that could not simply stem from the spectral overlap between the target and masker. Here, we replicate and extend this finding to the dichotic listening situation. Therefore, it seems that the linguistic nature of a babble noise impacts speech perception, even in the absence of spectral overlap between target and masker at the peripheral level. Further studies investigating dyslexic children's consonant identification in N-talker babble as a function of N might shed light on the contradictory results observed regarding IM contribution to their difficulties in perceiving speech in noise.

Interestingly, dyslexic children performed worse than age-matched controls even in the predictable condition, which was expected to reduce uncertainty related to target lateralization. This observation is in line with previous findings of poor use of auditory lateralization cues in dyslexic children (Smith and Griffiths, 1987; Calcus *et al.*, 2015).

However, our results show that dyslexic children were only impaired when compared to age-matched controls, not when compared to younger, reading level-matched controls. There are two potential explanations for this finding. First, dyslexic children could experience a maturational delay in the ability to understand speech presented in IM situations, which develops over time in normal-reading children (Wightman *et al.*, 2010). Consistently, dyslexic adults performed similar to controls in a dichotic word identification task presented in SSN, eSSN, and babble noise backgrounds and using a material similar to the present one (Dole *et al.*, 2012). A second (not necessarily incompatible) explanation is that reading acquisition could by itself strengthen auditory perception, as orthographic knowledge is known to influence performance in auditory tasks, especially in noise (Pattamadilok *et al.*, 2011). Whatever the explanation, the absence of deficit when compared to reading level-matched children, together with the absence of any correlation between dyslexics' performance in noise and either reading or metaphonological abilities, are inconsistent with previous studies claiming that auditory perception in noise is a core difficulty, inherent to dyslexia (Ziegler *et al.*, 2009; Calcus *et al.*, 2015). Yet, the sample used in the current study is quite small ($n = 30$). Investigating the impact of the specific nature of the background noise on speech intelligibility in larger samples would allow specifying the nature of the relationship between auditory perception deficit and reading disorder.

In conclusion, refining the methodology used to investigate speech perception in noise is necessary in order to specify the nature of the difficulty encountered by dyslexic

children. Indeed, investigating IM using speech material, our results point to a cognitive contribution to the speech-in-noise perception deficit observed in dyslexic children.

Acknowledgments

References and links

Bogliotti, C., Serniclaes, W., Messaoud-Galusi, S., and Sprenger-Charolles, L. (**2008**). "Discrimination of speech sounds by children with dyslexia: Comparisons with chronological age and reading level controls," J. Exp. Child Psychol. **101**, 137–155.

Brungart, D., Simpson, B., Ericson, M., and Scott, K. (**2001**). "Informational and Energetic masking effects in the perception of multiple simultaneous talkers," J. Acoust. Soc. Am. **110**, 2527–2538.

Calcus, A., Colin, C., Deltenre, P., and Kolinsky, R. (**2015**). "Informational masking of complex tones in dyslexic children," Neurosci. Lett. **584**, 71–76.

Chandrasekaran, B., Hornickel, J., Skoe, E., Nicol, T., and Kraus, N. (**2009**). "Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: Implications for developmental Dyslexia," Neuron **64**, 311–319.

Dole, M., Hoen, M., and Meunier, F. (**2012**). "Speech-in-noise perception deficit in adults with dyslexia: Effects of background type and listening configuration," Neuropsychologia **50**, 1543–1552.

Durlach, N., Mason, C., Kidd, G., Arbogast, T., Colburn, S., and Shinn-Cunningham, B. (**2003**). "Note on informational masking," J. Acoust. Soc. Am. **113**, 2984–2987.

Hazan, V., Messaoud-Galusi, S., and Rosen, S. (**2013**). "The effect of talker and intonation variability on speech perception in noise in children with dyslexia," J. Speech Lang. Hear. Res. **56**, 44–62.

Messaoud-Galusi, S., Hazan, V., and 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.

Miller, G., and Niceley, P. (**1955**). "An analysis of perception confusions among some English consonants," J. Acoust. Soc. Am. **27**, 338–352.

Mousty, P., and Leybaert, J. (**1999**). "Evaluation des habiletés de lecture et d'orthographe au moyen de BELEC: Données longitudinales auprès d'enfants francophones testés en 2ème et 4ème années" ("Evaluation of reading and orthographic abilities using BELEC: Longitudinal data from French speaking children in 2nd and 4th Grade"), Revue Européenne De Psychologie Appliquée **49**, 325–342.

Neff, D., and Green, D. (**1987**). "Masking produced by spectral uncertainty with multicomponent maskers," Percept. Psychophys. **41**, 409–415.

Pattamadilok, C., Morais, J., and Kolinsky, R. (**2011**). "Naming in noise: The contribution of orthographic knowledge to speech repetition," Frontiers Psychol. **2**, 1–12.

Poelmans, H., Luts, H., Vandermosten, M., Boets, B., Ghesquière, P., and Wouters, J. (**2011**). "Reduced sensitivity to slow-rate dynamic auditory information in children with dyslexia," Res. Dev. Disabilities **32**, 2810–2819.

Shinn-Cunningham, B. G. (**2008**). "Object-based auditory and visual attention," Trends Cognit. Sci. **12**, 182–186.

Simpson, S., and Cooke, M. (**2005**). "Consonant identification in N-talker babble is a nonmonotonic function of N," J. Acoust. Soc. Am. **118**, 2775–2778.

Smith, K., and Griffiths, P. (**1987**). "Defective lateralized attention for non-verbal sounds in developmental dyslexia," Neuropsychologia **25**, 259–268.

Snowling, M. (**2000**). "The definition of dyslexia," in *Dyslexia*, 2nd ed. (Blackwell, Oxford), pp. 14–29.

Stone, M., Füllgrabe, C., and Moore, B. (**2012**). "Notionally steady background noise acts primarily as a modulation masker of speech," J. Acoust. Soc. Am. **132**, 317–326.

Vellutino, F., Fletcher, J., Snowling, M., and Scanlon, D. (**2004**). "Specific reading disability (dyslexia): What have we learned in the past four decades?," J. Child Psych. Psych. **45**, 2–40.

Wightman, F., Kistler, D., and O'Bryan, A. (**2010**). "Individual differences and age effects in a dichotic informational masking paradigm," J. Acoust. Soc. Am. **128**, 270–279.

Zeng, F.-G., Nie, K., Stickney, G., Kong, Y.-Y., Vongphoe, M., Bhargave, A., Wei, C., and Cao, K. (**2005**). "Speech recognition with amplitude and frequency modulations," Proc. Natl. Acad. Sci. **102**, 2293–2298.

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