



# The Role of Slow Speech Amplitude Envelope for Speech Processing and Reading Development

Paula Ríos-López<sup>1\*</sup>, Monika T. Molnar<sup>1</sup>, Mikel Lizarazu<sup>2</sup> and Marie Lallier<sup>1</sup>

<sup>1</sup> Developmental Language Disorders, Basque Center on Cognition, Brain and Language, Donostia-San Sebastián, Spain,

<sup>2</sup> Laboratoire de Sciences Cognitives et Psycholinguistique, Centre National de la Recherche Scientifique, Ecole Normale Supérieure, Paris, France

## OPEN ACCESS

### Edited by:

Jerker Rönnerberg,  
Linköping University, Sweden

### Reviewed by:

Chloe Marshall,  
UCL Institute of Education,  
United Kingdom  
Mariapaola D'Imperio,  
Aix-Marseille University, France

### \*Correspondence:

Paula Ríos-López  
p.rios@bcbl.eu

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Psychology

**Received:** 26 April 2017

**Accepted:** 18 August 2017

**Published:** 31 August 2017

### Citation:

Ríos-López P, Molnar MT, Lizarazu M  
and Lallier M (2017) The Role of Slow  
Speech Amplitude Envelope for  
Speech Processing and Reading  
Development. *Front. Psychol.* 8:1497.  
doi: 10.3389/fpsyg.2017.01497

This study examined the putative link between the entrainment to the slow rhythmic structure of speech, speech intelligibility and reading by means of a behavioral paradigm. Two groups of 20 children (Grades 2 and 5) were asked to recall a pseudoword embedded in sentences presented either in quiet or noisy listening conditions. Half of the sentences were primed with their syllabic and prosodic amplitude envelope to determine whether a boost in auditory entrainment to these speech features enhanced pseudoword intelligibility. Priming improved pseudoword recall performance only for the older children both in a quiet and a noisy listening environment, and such benefit from the prime correlated with reading skills and pseudoword recall. Our results support the role of syllabic and prosodic tracking of speech in reading development.

**Keywords:** reading, amplitude envelope, phonological processing, speech-in-noise, auditory entrainment, language development

## INTRODUCTION

The amplitude envelope of speech contains crucial information about prosodic and stress patterns, speech rate, tonal contrasts and intonational information necessary for successful speech perception. Specifically, multiple studies have shown that the slow components of the speech amplitude envelope (below 8 Hz) – that code for prosodic and syllabic information (Houtgast and Steeneken, 1985) – contribute significantly to the perception of the rhythmic structure of speech (Rosen, 1992; Hickok and Poeppel, 2007; Abrams et al., 2008; Giraud and Poeppel, 2012).

Pioneering behavioral evidence on the role of the slow temporal fluctuations of speech for successful speech perception has received notable neurophysiological support during the last decade. In particular, numerous studies have shown that the oscillatory activity in the brain “phase-locks” to the temporal structure of auditory speech stimuli (Suppes et al., 1997, 1998; Ahissar et al., 2001; Abrams et al., 2008; Aiken and Picton, 2008; Nourski et al., 2009; Peelle and Davis, 2012; Bourguignon et al., 2013; Peelle et al., 2013; Molinaro et al., 2016). One of the reasons why experiments have proliferated in this field concerns the fact that endogenous brain oscillatory activity which coincides with typical frequencies of salient speech features (i.e., gamma, theta and delta, with phonemes, syllables, and stress, respectively) could represent a good candidate mechanism that would support speech parsing (Ghitza, 2012; Poeppel, 2014; Leong and Goswami, 2015): phonemes appear at rates of approximately 80 ms (gamma band, >25 Hz), syllables at approximately 200 ms (theta band, 4–8 Hz), and stressed syllables at around 500 ms (delta band <4Hz; Ghitza and Greenberg, 2009; Goswami, 2011). Based on this phenomenon, several studies have described the properties of neural oscillatory mechanisms in relation to the human ability to

encode and segment speech (e.g., Poeppel, 2003; Giraud et al., 2007; Goswami, 2011; Giraud and Poeppel, 2012; Gross et al., 2013). For example, according to the “asymmetric sampling in time” hypothesis (Poeppel, 2003; Giraud and Poeppel, 2012), tracking the speech amplitude envelope at prosodic and syllabic rates would allow the brain to rely on some acoustic landmarks on which to anchor “neural entrainment,” i.e., synchronization of the brain oscillatory activity to external (e.g., auditory) signals. Complementary hypotheses suggest that neural oscillatory mechanisms would allow neuronal responses to be amplified and highly excitable when critical phonological, prosodic and syllabic events occur in speech streams (Lakatos et al., 2008; Schroeder and Lakatos, 2009).

Supporting these hypotheses, several experiments found a positive correlation between speech intelligibility and neural entrainment to the prosodic and syllabic properties of the auditory speech stimuli. Relevant to the current study, speech comprehension is accompanied by enhanced phased-locking values to slow amplitude speech modulations in the auditory cortex (Peelle and Davis, 2012; Peelle et al., 2013). The importance of perceiving the structure of the slow amplitude envelope of speech may be even more significant when speech intelligibility is challenged by competing background noises (which is the most common everyday listening situation). Accordingly, the role of such an amplitude-tracking mechanism during speech perception in noise has been investigated by several studies (Zion Golumbic et al., 2012, 2013; Ding and Simon, 2014). Specifically, studies on selective attention suggest that neural phase locking to the slow speech amplitude envelope of an auditory attended stimulus may help to ignore other simultaneous auditory stimuli by enhancing the selection of information to be attended and reducing the allocation of attentional resources to information that needs to be ignored (Schroeder and Lakatos, 2009; Kerlin et al., 2010; Ding and Simon, 2012; Zion Golumbic et al., 2012).

In a different line of research, perception of the slow amplitude envelope of speech has also been linked with reading abilities. Ample evidence supports the idea that atypical perception of the amplitude envelope of speech might lead to an abnormal acquisition of phonological representations, which would in turn lead to reading difficulties such as developmental dyslexia (Goswami et al., 2002, 2010). For example, Goswami et al. (2010) showed that processing of amplitude envelope structure, — specifically rise time sensibility— predicted phonological awareness skills, as measured via a rhyming task.

In line with the hypotheses on the role of oscillatory mechanisms for speech perception mentioned above, it has been proposed that phonological disorders underlying reading difficulties could stem from an atypical auditory sampling of temporal speech features (e.g., Goswami, 2011; Lehongre et al., 2013). The “temporal sampling framework” hypothesis suggests that reading difficulties in dyslexic individuals would stem from an atypical (asynchronous) tracking of the speech amplitude envelope corresponding to prosodic information, such as syllabic stress, in particular (Goswami, 2011; Goswami and Leong, 2013). In this framework, atypical temporal sampling of syllabic stress is thought to hamper the development of phonological

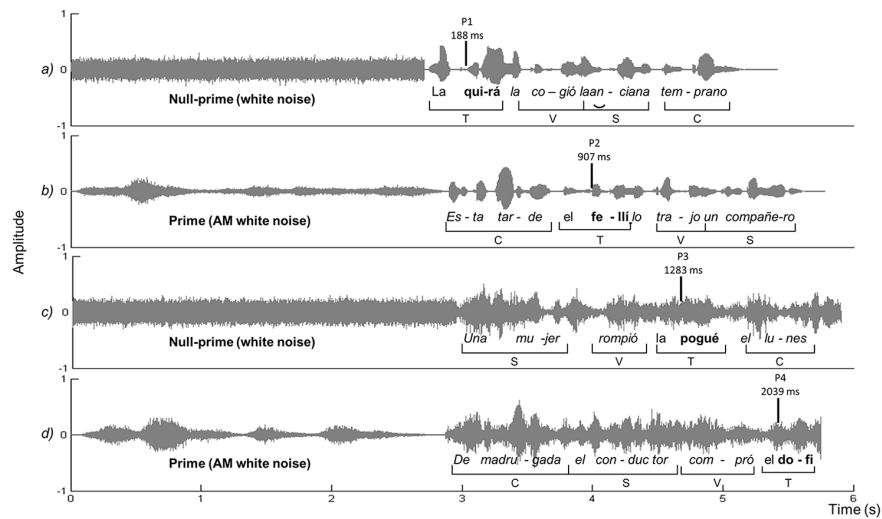
skills, and in turn, reading acquisition. Accordingly, there is consistent evidence showing impaired synchronization to the speech amplitude envelope at low frequencies in dyslexic children and adults (Hämäläinen et al., 2005, 2012; Corriveau et al., 2007; Abrams et al., 2009; Goswami et al., 2010; Power et al., 2013). Hämäläinen et al. (2012) showed that dyslexic adults exhibited weaker neural phase locking to noises modulated in amplitude at 2 Hz compared to controls. The results of Hämäläinen et al. (2012) suggest that normal reading acquisition may depend on the adequate neural and attentional synchronization to prosodic and syllabic speech events. Interestingly, the temporal sampling framework, Goswami (2011; Goswami et al., 2014) can explain why difficulties in integrating speech information at low rates could also cause problems to integrate speech units that oscillate at higher frequencies, such as phonemes. This is a critical point in Goswami’s theory, since it offers an explanation why impaired representations and manipulation of phonemes strongly predict reading disorders across alphabetic languages (Hulme et al., 2005; Ziegler et al., 2010; Caravolas et al., 2012). Neural oscillations at low and high frequency bands are not independent. In particular, cortical oscillations reflect a “nesting” behavior such that the power of the brain responses in high frequency bands couples to the phase of lower frequencies generating a coupling (synchrony) between various neural oscillatory rates (Lakatos et al., 2005; Canolty et al., 2006; Gross et al., 2013). In other words, a precise temporal allocation of attention over prosodic and syllabic temporal speech features (i.e., subtended by neural entrainment) would enhance attentional focusing on phonemic information that falls within the attentional window (Lallier et al., 2017). Hence, the nesting of neural oscillators can explain why fast phonemic integration might depend on the efficient coding of slower speech rates (Ghitza, 2011; Giraud and Poeppel, 2012).

In addition to auditory oscillatory processing deficits, children and adults with reading difficulties also exhibit pronounced difficulties in speech-in-noise tasks (Wible et al., 2002; Bradlow et al., 2003; Ziegler et al., 2009; Boets et al., 2011; cf. Hazan et al., 2009). Such difficulties in dyslexia may arise from impoverished intelligibility of phonemic information since under noise circumstances, the saturation of the signal at the phonemic level is high and listeners cannot identify phonemic features as efficiently as in quiet conditions (i.e., “informational masking” in speech-in-noise, Brungart et al., 2001). Hence, it is possible that speech-in-noise difficulties associated with reading deficits also result from impaired auditory entrainment to speech amplitude envelope at prosodic and syllabic slow rates.

## The Current Study

Taking into account the above-described evidence on the relation between slow-amplitude speech envelope processing and spoken language intelligibility, in addition to the putative link of these aspects with reading, the aim of this study was to test the link between auditory entrainment to speech, understood as the ability to synchronize to the (quasi)rhythmic modulations in the speech signal, and reading skills in typically developing children through a behavioral paradigm.

In this experiment, children of Grades 2 and 5 listened to Spanish sentences presented in either quiet or noisy



**FIGURE 1** | This figure shows the audio signal of four experimental sentences in the temporal domain, with time in seconds in the x-axis and amplitude in the y-axis. Because four different experimental sentences are presented, slight sentence duration differences are present. The mid-left part of the graph shows the prime (Null-prime vs. Prime) and the mid-right part the experimental sentence (including the target pseudoword). Target pseudowords are highlighted in bold. Sentence (a) corresponds to the Quiet null-prime condition; sentence (b) is an example of the Quiet prime condition. Sentences (c) and (d) correspond to the Noise listening condition (six-talker babble) in the Null-prime condition [sentence (c)] and in the Prime condition [sentence (d)]. AM, amplitude modulated.

listening conditions, before they had to recall an embedded target pseudoword from the sentences (see **Figure 1** for a graphical depiction of the paradigm). Importantly, each sentence was primed with its own slow amplitude envelope that was constructed of white noise modulated with the given sentence frequencies below 8 Hz (adapted from Peelle and Davis, 2012; for further details on the stimuli, see the Materials and Methods section). In the experiment, first the children heard the slow amplitude envelope of the target sentence as prime (that contained no phonemic information, only prosodic and rhythmic information), and then the children heard the target sentence natural production (that contained all the acoustic-linguistic information of speech, including the phonetic and prosodic information). This way, only the sentence slow amplitude envelope was repeated across the prime and the target sentence.

The aim of this design was to test whether it is possible to boost phonemic encoding and the target pseudoword intelligibility in the target sentence by enhancing auditory entrainment to the slow rhythmic structure of the sentence (through priming and/or repeating the sentence amplitude envelope). Because the auditory cortical activity couples with the temporal envelope of speech (Suppes et al., 1997, 1998; Ahissar et al., 2001; Abrams et al., 2008; Aiken and Picton, 2008; Nourski et al., 2009; Peelle and Davis, 2012; Bourguignon et al., 2013; Peelle et al., 2013; Molinaro et al., 2016), we predicted the sentence amplitude envelope prime to elicit auditory entrainment to the (quasi)rhythmic properties of the sentence before listening to the spectrally complete sentence. Specifically, we predicted that such preparation/priming would enhance the processing of the same subsequent rhythmic structure, and, in turn, its intelligibility through better phonemic encoding. Indeed, recent research in the field of cross-domain effects of music on speech

has shown that priming speech with music rhythms that match speech prosodic features improves phonological processing, as measured via pseudoword recall (Cason and Schön, 2012), and that rhythmic priming also contributes to better performance in other language related tasks, such as in a grammaticality judgment task (Przybylski et al., 2013).

We decided to use the low-pass filtered (<8 Hz) amplitude envelope of the sentences rather than a periodic rhythmic stimulation at a specific frequency for two main reasons: (1) by priming the target sentence with its slow amplitude envelope below 8 Hz, we were able to prime both the prosodic and syllabic rhythmic structures of the subsequent sentence in order to maximize the potential beneficial effect of the prime; (2) because we know that entrainment is triggered by statistical regularities different from periodic signals (Obleser et al., 2012; Goswami and Leong, 2013), such as strong-weak syllable stress patterns (Rothermich et al., 2012), we considered that the aperiodic statistical regularities such as the ones contained in our prime would be good candidates to elicit entrainment.

If slow speech amplitude envelope tracking is critical for speech intelligibility, priming the sentence with its slow amplitude envelope (<8 Hz) should foster auditory entrainment to both prosodic (0–4 Hz) and syllabic (4–8 Hz) speech rhythms, and improve pseudoword recall (Lallier et al., 2017).

We also presented the target sentences in noise. We predicted that if speech amplitude envelope tracking at prosodic and syllabic rates plays an essential role in the segregation of a target signal from a noisy environment, the primes should also increase the target sentence intelligibility presented in babble noise (“Cocktail Party effect”; Cherry, 1953). Since the speech amplitude envelope at prosodic, syllabic and phonemic rates is strongly masked by this type of noise (Brungart et al.,

2001), we expected performance to be considerably impoverished in the null-prime condition, as opposed to in the prime condition. It was assumed that priming noisy sentences with their prosodic- and syllabic-rate structure would enhance auditory entrainment to the rhythmic structure of the sentence helping children to (selectively) attend to prosodic and syllabic amplitude fluctuations in the target sentence and facilitating in turn phonemic processing at a faster rate. In the prime condition, pseudoword recall performance was therefore expected to improve compared to the null-prime condition.

Based on current theories on reading disorders and dyslexia (Goswami, 2011), we finally hypothesized that children with low reading skills would be less sensitive to the slow speech amplitude envelope in null-primed sentences. Therefore, we expected low reading skills to be linked to larger benefit from the speech amplitude envelope prime on their pseudoword recall. Indeed, children with good reading skills should be overall more sensitive than children with low reading skills to prosodic and syllabic speech features and would not be greatly helped from the prime.

In the line of the present work, several previous studies have tested the predictions of neurophysiological hypotheses, namely of multi-temporal resolution models as the “asymmetric sampling in time” model (Poeppel, 2003), using behavioral paradigms with filtered speech (e.g., Chait et al., 2015; Goswami et al., 2016). For example, Goswami et al. (2016) used low-pass filtered speech (at <4 Hz) in a speech perception behavioral task to support the hypothesis that dyslexic children have poor neural synchronization to slow speech amplitude modulations. The authors showed impoverished perception of low-passed filtered speech (i.e., speech where only prosodic information was available) that they interpreted as behavioral evidence for deficits in neural speech sampling in dyslexia.

## MATERIALS AND METHODS

### Participants

Forty children (17 males; mean age 9 years 3 months,  $SD = 20.6$  months) were recruited for the experiment in a school in Donostia-San Sebastián (Basque Country, Spain). The first group consisted of 20 children attending the 2nd grade (9 males; mean age 7 years 7 months,  $SD = 2.2$  months). The second group consisted of 20 children attending the 5th grade (8 males; mean age 11 years,  $SD = 4.0$  months). Four of the children in the group of 5th graders were left-handed.

Donostia-San Sebastián is a bilingual community where both Basque and Spanish coexist as official languages, and therefore parents were asked to fill out a questionnaire about the linguistic profile of their children. Following the analysis of the responses to the questionnaire, all children were classified as bilingual of Spanish and Basque. We used a measure of general exposure to the specific languages (in %) to determine if the children were Basque or Spanish dominant. Children with percentage of exposure to a language (Basque or Spanish) of more than 70% were considered dominant in that language. This analysis resulted in five children in the group of 2nd graders and one child in the group of 5th graders who were Basque-dominant. None of

the children was diagnosed with any hearing disability, language-based or developmental disorder of any other kind, and it was stated that they suffered no ear infections at the time of testing. All children were recruited from a private school in Donostia-San Sebastián and tested under signed parental consent.

## Speech Perception Task

### Stimuli

#### Target pseudowords

Target items were pseudowords with a low-density lexical neighborhood to avoid the influence of lexical information and require the participants to rely more strongly on sublexical phonological processing (Peelle and Davis, 2012; Peelle et al., 2013). The neighborhood density was calculated using the Levenshtein distance calculator of Yarkoni et al. (2008; mean neighborhood density 8.84).

To avoid any effects on target recall performance that could be due to the repetition of specific phonological and acoustic parameters across targets, we controlled for their length, stress pattern and phonemic structure. The target pseudowords were two- and three-syllable long items, and all syllables followed a consonant-vowel pattern (CVCV or CVCVCV). Half of the items ( $n = 40$ ) were stressed in the first syllable and the other half ( $n = 40$ ) in the last syllable. We counterbalanced the number of voiced and unvoiced consonants and vowels to avoid confounding variables due to sound complexity (difference in rise time, duration, etc.). Repetition of phonemes within targets was avoided. Finally for the properties of the pseudowords, only high-frequency bigrams were used (see Edwards et al., 2004 for manipulation of bigram frequency and its interaction with vocabulary size). Mean bigram frequency of the syllables was calculated with the tool *Espal* and was found to be overall high (Sample Mean = 7656,  $SD = 3948$ ; database Mean = 5179; Duchon et al., 2013). The pseudoword was always preceded by a definite article, with gender counterbalanced across stimuli (*el* or *la*), and it was embedded in a Spanish sentence as the direct object of the verb (see Carrier Sentences for the different constituents composing the sentence).

#### Carrier sentences

Twenty different sentences were created. Sentences consisted of six to eight words ( $M = 7.0$  words,  $SD = 0.6$ ) and varied from 1.97 to 3.45 s ( $M = 2.8$  s,  $SD = 0.3$ ) in duration including four grammatical constituents. They were recorded by a female native speaker of Spanish in a sound-proof recording room. The sentences were not subject to digital manipulation, and the speaker read the pseudoword naturally within the sentence. The reader was blind to the purpose of the experiment and was instructed to read at a comfortable pace with a normal prosody. The sample frequency was fixed at 44.1 kHz.

Four grammatical constituents composed each sentence: (i) a subject phrase including a human agent whose gender was counterbalanced across sentences (e.g., *la abuela* [the grandmother]), (ii) a time complement which consisted of a high-frequency temporal adverb (e.g., *ayer* [yesterday]) or adverbial phrase (e.g., *por la mañana* [in the morning]), (iii) a high frequency transitive verb (log count  $M = 4.1$ ,  $MED = 4.0$ ,



$SD = 0.4$ ) conjugated in the past simple, (iv) an object phrase which head was the target pseudoword (e.g., *la befu*, see Target Pseudowords for further details).

In order to avoid expectancy effects (i.e., children paying attention only to the target pseudoword, since the task was only to recall this element), we placed the target pseudoword in four different positions across our carrier sentences. Word order is highly flexible in Spanish, and this manipulation resulted in four different Spanish legal grammatical sentence structures (see **Table 1** for details on word order and time of appearance of the pseudoword within the carrier sentences).

### Primes

The primes were created via Matlab® (Mathworks, Natick, MA, United States) following the procedure for noise vocoding described in Peelle and Davis (2012). The audio signal from each sentence was divided into four logarithmic spaced channels. We decided to divide the signal into four channels to obtain an unintelligible, although still not excessively unpleasant sound for the children. The spectrum of each channel was extracted and a Hilbert transform was applied to extract the amplitude envelope of each signal. The signals were low-passed filtered at 8 Hz and used to modulate white noise in the same frequency bands. Finally, the channels were recombined to create the final amplitude envelope signal. The result was a modulated white noise which contained minimal phonemic information and matched the prosodic and syllabic amplitude envelope and the duration of the original sentence (see **Figure 1**). The prime was placed before the sentence with inter-stimulus interval (ISI) of 0 ms, since our intention was to create a continuous stimulus and gaps would encourage the discontinuity of attentional deployment.

For the null-prime conditions, we created a stream of random white noise that matched the original sentence in duration but contained no amplitude envelope information (steady white noise). The null-prime was also placed before the sentence with an ISI of 0 ms.

### Noise: 6-talker babble

The babble-noise was constructed following the procedure used in Dole et al. (2012). Six speakers (three males, three females) were recorded in a soundproof room whilst reading passages of Spanish newspapers. From these individual six recordings, silences of more than 1 s were removed. Fragments containing pronunciation errors or exaggerated prosody were also discarded. We then applied a noise reduction to eliminate artifact interference for each of the six individual tracks, and, finally, mixed the tracks to create the 6-talker babble. The babble noise stimuli were then superimposed to half of the

carrier sentences ( $n = 80$ ) using Matlab® scripts at a SNR of 7 dB.

### Procedure

Participants were sat in a quiet dim-lit room and listened to the sentences one by one through headphones (Sennheiser) and delivered via the Presentation® software. Stimuli were normalized at 80 dB and presented binaurally. The children were explicitly instructed to listen to the whole auditory stimulus (prime + sentence), but they were not informed about the aim of the experiment or the role of the prime. After the sentence, the experimenter asked a specific question that required the child to recall and produce the target pseudoword (e.g., Sentence: *Ayer la befu la tiró la madre* [*Yesterday the mother dropped the befu*]; Question: *¿Qué tiró la madre ayer?* [*What did the mother drop yesterday?*]; Response: *la befu*). Participants could respond at their own pace and the next sentence was played only after the response was given. Responses were recorded by a mobile recording device (Zoom H2) and coded by the experimenter.

The same 20 carrier sentences were repeated across the four prime\*noise conditions in a random order. In that way, all carrier sentences were presented in a prime and null-prime condition within each listening condition (quiet and noise). Importantly, although the carrier sentences were the same across conditions, participants never listened to the same target pseudoword twice to avoid repetition effects on target perception and recall (total of 80 pseudowords). To avoid effects of the intrinsic characteristics of pseudowords (i.e., recall performance due to the characteristics of the pseudowords and not its time frame of appearance), two subgroups were created within each age group. The two subgroups listened to the same pseudowords and to the same carrier sentences, but recombined in a different way, i.e., pseudowords were not embedded in the same carrier sentences for both groups, and hence they listened to the target pseudowords in different conditions.

Eight practice sentences were presented prior to the experimental task. Breaks were made when necessary. The total duration of the experiment was approximately 15 min.

## Neuropsychological Screening

### Non-verbal IQ

Non-verbal IQ was measured with the Matrix reasoning task of the WISC-IV. Direct scores were converted into scalar scores.

### Phonological and Reading Tasks

#### *Pseudoword repetition*

This type of task has been extensively used to measure phonological short-term memory (e.g., Dollaghan and Campbell,

**TABLE 1** | Word order and mean time of appearance of the Target (pseudoword) within the carrier sentences (P, position).

	Position of constituents within the sentence				Mean time appearance of the Target in ms (SD)
Target in P1	<b>Target</b>	Verb	Subject	Complement	188 (44.5)
Target in P2	Complement	<b>Target</b>	Verb	Subject	907 (168.3)
Target in P3	Subject	Verb	<b>Target</b>	Complement	1283 (137.5)
Target in P4	Complement	Subject	Verb	<b>Target</b>	2039 (211.7)

1998; Lallier et al., 2012). Participants were asked to repeat pseudowords varying in length (3 items of 2 syllables, 3 of 3 syllables, 3 of 4 syllables, and 3 of 5 syllables; total of 12). The syllable structure always followed a consonant-vowel (CV) pattern. The items were delivered through headphones one by one, and the children were asked to recall them. The final score corresponded to the total number of correct responses produced by the child (maximum = 12).

### Phonemic deletion

This type of task has been extensively used to measure phonemic awareness (e.g., Stanovich et al., 1984). After a pseudoword was delivered through headphones, the participants were asked to repeat the entire item. Then, they had to repeat the item again but this time they were asked to omit the first sound. The task consisted of a total of 12 bi-syllabic items delivered in a random order (e.g., /timu/ → /imu/). The final score corresponded to the total number of correct responses produced by the child (maximum score deletion based on the repetition of the child = 12).

### Text reading

Participants were asked to read a short passage from the novel “The little prince” (“El principito,” written by Saint Exupéry) translated into Spanish (taken from Lallier et al., 2014) and consisting of 103 words and 9 lines. Participants were instructed to read as fast as possible minimizing the number of errors. Time (in seconds) and number of errors were recorded. Number of errors and total reading time were calculated as two separated scores.

### Single word and pseudoword reading

Participants read single words ( $n = 30$ ) and pseudowords ( $n = 30$ ) taken from PROLEC-R (battery of evaluation of reading processes in Spanish; Cuetos et al., 2007). Participants were instructed to read as fast and as well as possible. Time and number of errors were recorded. Number of errors and total reading time were calculated as two separate scores.

## DATA ANALYSIS

To examine between-group differences in non-verbal IQ, reading and phonological skills, two-tailed  $t$ -tests were computed with grade as between-subject factor and Matrices (scalar scores), reading skills (word reading time, word reading errors, pseudoword reading time, pseudoword reading errors), and phonological skills (phoneme deletion scores, pseudoword repetition scores) as dependent variables.

For the speech perception task, the analysis was carried out fitting linear mixed-effect models (Baayen et al., 2008) with accuracy of target production (i.e., correct repetition of the whole target) as a dependent variable, and grade (two levels: Grades 2 and 5), listening condition (two levels: quiet and noise) and prime (two levels: null-prime and prime) as predictors. The complexity of the random-effects structure was increased in a series of models (cf. Barr et al., 2013), leading to a maximal converging model that contained all

random effects that improved significantly the residual variance captured by the model. The *lme4* package (Bates et al., 2013) contained in R (R Development Core Team, 2011) was used to perform the mixed-effects models analysis. For analysis of accuracy measures, this package uses the inverse link function defined in the GLM family, through which the linear predictor is related to the conditional mean of the response. The output is hence expressed in *logits*. For clarity reasons, in the text we report inverse transformation from *logits*.

The random effects structure included items and subjects, and we added an additional random effect for target positions (1, 2, 3, 4). Although the four different sentence structures were grammatically legal, Spanish is a predominantly SVO language, and hence we expected that position of the target pseudoword could affect pseudoword recall. Regarding the interactions between the different fixed effects, we followed the same procedure, i.e., starting from the simplest model with no interactions and adding interactions that improved the fit of the data.

Before the recall scores obtained on the speech perception task were submitted to statistical analyses, all the oral responses of the children were double-checked by another native speaker of Spanish who was blind to the experimental conditions. The original coding of the experimenter and the verification coding matched for 95% of the items. A final decision for the items in conflict was reached upon agreement of both judges.

To perform correlation analyses, we calculated the benefit that the children obtained from the prime on pseudoword recall by subtracting the mean recall accuracy in the null-prime condition from the mean recall accuracy in the prime condition in both listening conditions separately (i.e., in quiet and in noise). Two-tailed partial correlations between the prime benefit values and reading as well as phonological skills were computed for both grades separately controlling for non-verbal IQ and chronological age.

## RESULTS

### Neuropsychological Screening

Regarding text reading time [ $t(37) = 5.39$ ;  $p < 0.001$ ], 5th graders were significantly faster ( $M = 51.3$  s;  $SD = 16.9$ ) than 2nd graders ( $M = 102.4$  s;  $SD = 38.7$ ; see **Table 2**). No other difference between the two grades was found on non-verbal IQ, phoneme deletion, pseudoword repetition, text reading errors, word reading time, word reading errors, pseudoword reading time and pseudoword reading errors ( $ps > 0.05$ ; see **Table 2**). Four of the children in Grade 5 were left-handed. We did not perform additional analyses to rule out the possibility that this could be a contributing factor to the group differences we reported, since hand-dominance does not seem to play a significant role in the distribution of cerebral asymmetries for processing temporal information of acoustic signals (Abrams et al., 2009). Language dominance (Basque vs. Spanish) did not correlate with any of the measures assessed in this study.

**TABLE 2** | Neuropsychological screening.

	Grade 2 ( <i>n</i> = 19) <i>M</i> ( <i>SD</i> )	Grade 5 ( <i>n</i> = 20) <i>M</i> ( <i>SD</i> )	Group difference <i>t</i> -values (2-tailed), <i>p</i> -values	
Chronological age (months)	91.9 (2.2)	132 (4)	−39.2	<0.001
Non-verbal IQ (WISC matrix subtest)	13.05 (3.1)	11.7 (2.5)	1.6	0.13
Reading skills				
Word reading – Number of errors	0.5 (0.77)	1.2 (1.5)	−1.7	0.1
Word reading – Time (s)	30.7 (12.93)	35.7 (10.6)	−1.33	0.2
Pseudoword reading – Number of errors	1.8 (1.80)	1.4 (1.2)	1.00	0.3
Pseudoword reading – Time (s)	50.4 (17.02)	53.6 (15.7)	−0.6	0.6
Text reading – Number of errors	2.4 (2.09)	2.3 (2.0)	0.2	0.9
Text reading – Time (s)	102.4 (38.7)	51.3 (16.9)	5.4	<0.001
Phonological skills				
Pseudoword repetition – Accuracy/12	10.2 (1.4)	10.4 (1.3)	−0.6	0.6
Phonemic deletion – Accuracy/12	9.2 (2.5)	8.7 (2.0)	0.9	0.4

## Speech Perception Task

Average accuracy of the two groups of children through the different conditions of the speech perception task is detailed in **Tables 3, 4**.

As mentioned above, the complexity of the model was increased gradually following the procedure described in Barr et al. (2013). Specifically, a first model that included random intercept for Subject was successively compared to subsequent mixed-effect models containing by-Subject, by-Item and by-Position random intercepts. Random slopes for the different effects were added thereafter to allow the slope of the fixed effects (Grade, Prime, and Listening condition) to fluctuate across subjects, items and positions. The limitation of the estimation algorithm led to a simplification of the random effect structure with a maximal converging model that included random intercepts by-Subject (*SD* = 0.67), by-Item (*SD* = 0.76) and by-Position (*SD* = 0.64), and random slopes by-Subject for the effect of Prime (*SD* = 0.63) and by-Position for the effect of Listening condition (*SD* = 0.63). Results of the likelihood ratio test and changes in the coefficients of the model were used to check if including additional random effects improved the fit of the model. For each comparison, we report the intercept, the estimated regression coefficients (Estimate,  $\beta$ ), the standard errors (SE), *z*- and *p*-values resulting from the mixed-effect model analysis.

**TABLE 3** | Mean accuracy (standard deviation in parenthesis) per Grade and Priming condition, averaged across participants for the Speech perception task.

	Grade 2 ( <i>n</i> = 20)	Grade 5 ( <i>n</i> = 20)
Null-Prime	0.51 (0.50)	0.62 (0.49)
Prime	0.52 (0.50)	0.72 (0.45)

**TABLE 4** | Mean accuracy (standard deviation in parenthesis) per Grade and Listening condition, averaged across participants for the Speech perception task.

	Grade 2 ( <i>n</i> = 20)	Grade 5 ( <i>n</i> = 20)
Quiet	0.76 (0.43)	0.91 (0.29)
Noise	0.27 (0.44)	0.43 (0.49)

A smaller probability of successful recall of the target pseudoword was found for the Noise (mean = 0.34, *SD* = 0.23) as compared to the Quiet (mean: 0.84; *SD*: 0.14) listening condition ( $\beta$  = 0.92; *SE* = 0.62; Wald's *z* = 5.19; *p* < 0.001). Regarding the predictor Grade, the probability of successful answer was greater for the 5th graders (mean = 0.67, *SD* = 0.22) as compared to the 2nd graders (mean = 0.52, *SD* = 0.25), but this difference was only marginally significant ( $\beta$  = 0.62; *SE* = 0.57; Wald's *z* = 1.81; *p* = 0.07). No main effect of prime was found in the maximally converging model (Null-prime mean = 0.56; *SD* = 0.25; Prime mean = 0.63; *SD* = 0.24;  $\beta$  = 0.47; *SE* = 0.60; Wald's *z* = −0.30; *p* = 0.77).

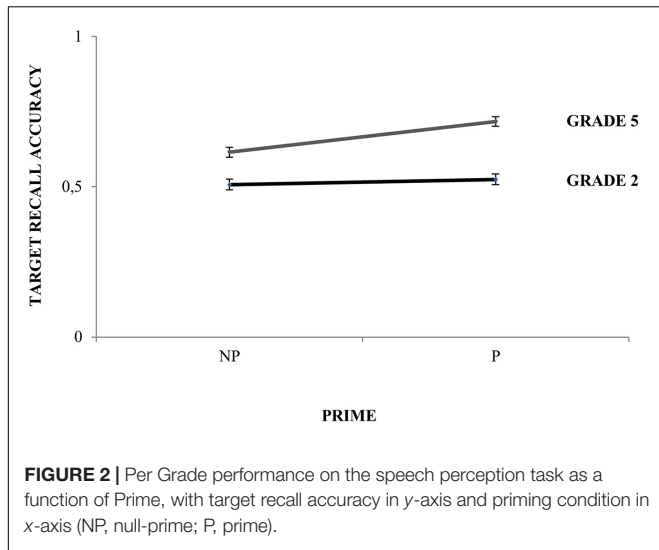
Regarding the interactions of the main predictors, the analysis revealed that Grade and Prime interacted significantly ( $\beta$  = 0.68; *SE* = 0.56; Wald's *z* = 2.96; *p* < 0.01), such that 5th graders (Null-prime mean = 0.62, Prime mean = 0.72) obtained a larger benefit from the prime as compared to 2nd graders (Null-prime mean = 0.51; Prime mean = 0.52; see **Figure 2** and **Table 3**). The interaction Listening condition by Grade was also significant ( $\beta$  = 0.63; *SE* = 0.55; Wald's *z* = 2.60; *p* < 0.01), revealing that performance was specially degraded in noise for the 2nd graders (Quiet mean = 0.76; Noise mean = 0.27) as compared to the 5th graders (Quiet mean = 0.91; Noise mean = 0.43; see **Table 4**).

Lastly for the fixed effects, the analysis revealed no significant Prime by Listening condition interaction ( $\beta$  = 0.67; *SE* = 0.64; Wald's *z* = 1.26; *p* = 0.21).

In light of the Listening condition by Grade interaction, we performed a follow-up analysis to examine the possibility that the significant by-Position random slope for the effect of Grade was driven by one of the groups only. While the slope and intercept for the 5th graders remained considerably constant across Listening conditions and Positions, the inclusion of position as a random-effect in the final model seemed to be reflecting only the performance pattern of the 2nd graders (see **Figure 3** and **Table 5**).

## Correlation Analyses

One child from Grade 2 was removed from the correlation analyses, since he had not completed some of the phonological and reading tasks.



**FIGURE 2 |** Per Grade performance on the speech perception task as a function of Prime, with target recall accuracy in y-axis and priming condition in x-axis (NP, null-prime; P, prime).

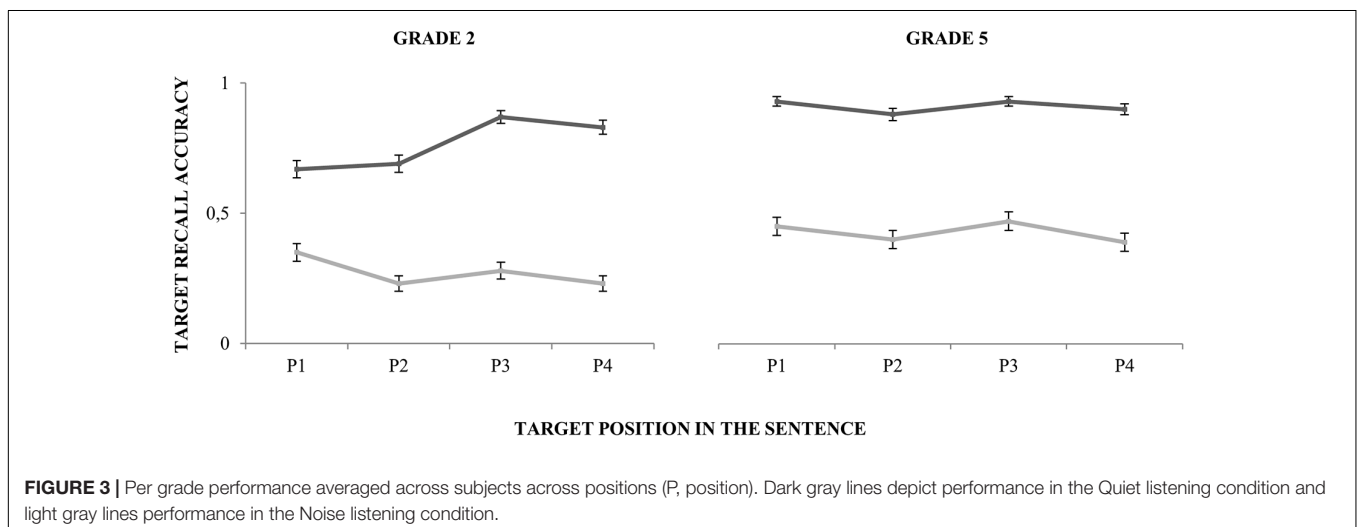
Within the group of the 2nd graders, no correlation was found between performance on the speech perception task and neither reading nor phonological skills (see **Table 6**). Regarding the 5th graders, a negative correlation between pseudoword repetition scores and the benefit obtained from the prime in the noisy conditions was found ( $r = -0.46, p < 0.05$ ), showing that the poorer the pseudoword repetition scores, the larger the prime benefit to recall the pseudoword presented in a noisy listening condition (**Figure 4**). Moreover, the slower the reading, the higher the benefit from the prime to recall pseudowords presented in a noisy listening condition (text reading:  $r = 0.53, p = 0.02$ ; word reading:  $r = 0.56, p = 0.01$ ; pseudoword reading:  $r = 0.52, p = 0.02$ ; see **Figure 4**).

## DISCUSSION

The present study assessed whether spoken language comprehension of typically developing school-aged children

can be enhanced by priming sentences with their own slow amplitude envelope only, in both quiet and noisy listening conditions. By priming sentences with their amplitude envelope low-pass filtered at 8 Hz (i.e., the speech amplitude envelope information was repeated across the prime and target sentence), we highlighted the impact of tracking the slow rhythmic structure of speech at prosodic and syllabic rates on speech intelligibility in Grade 5 children. In general, we observed that the processing of the sentences that were primed with their own speech envelope, as opposed to the sentences that were primed with non-modulated white noise (null-prime condition), were improved in children – as reflected by their better performance in pseudoword recall. We suggest that the speech amplitude envelope primes (without phonetic information) enhanced the processing of the speech amplitude envelope in the target sentences (via repetition of the speech amplitude envelope information), and this enhanced speech amplitude envelope processing enabled easier processing of the phonetic and lexical information in the target sentences. This study adds up to accumulating behavioral evidence (e.g., Chait et al., 2015; Goswami et al., 2016) supporting the predictions of multitemporal resolution neural models of speech (Zatorre and Belin, 2001; Boemio et al., 2005; Giraud et al., 2007; Obleser et al., 2008; Ghitza, 2011; Giraud and Poeppel, 2012; Luo and Poeppel, 2012; Saoud et al., 2012).

As predicted, our results revealed a main effect of listening condition, such that performance in pseudoword recall decreased significantly in the presence of the babble-noise. Interestingly, the prime led to a benefit in pseudoword recall across the quiet and the noise listening conditions similarly (absence of Listening condition by Prime interaction). However, the benefit from the prime seemed to be modulated by the age of the children. Group differences were observed such that the prime was overall more effective for the 5th graders than the 2nd graders. Nevertheless, we cannot discard that the effect of the prime for Grade 2 children was obscured by the high variability of this group’s performance, especially in the quiet listening condition (Quiet listening condition: Grade 2  $SD = 0.43$ ; Grade



**FIGURE 3 |** Per grade performance averaged across subjects across positions (P, position). Dark gray lines depict performance in the Quiet listening condition and light gray lines performance in the Noise listening condition.



5  $SD = 0.29$ ; Noise listening condition: Grade 2  $SD = 0.44$ ; Grade 5  $SD = 0.49$ ; see **Table 4**). Indeed, when considering individual prime benefits, eight out of the 20 children within Grade 2 showed an average effect of the prime greater than 5% across environments (12 children showed an effect of the prime larger than 5% in quiet). Hence, it could be that the variability in individual differences on the prime benefit could reflect different stages of development, which might be related to reading acquisition (see section *Relation between reading and entrainment to slow speech oscillations*). Related to this point, non-word repetition has been tightly related to vocabulary size (Edwards et al., 2004) and differences in vocabulary size could explain part of the differences between our groups. However, our paradigm was not designed to disentangle phonological short-term memory and vocabulary effects, and further research is needed to shed light on this relation.

Although our results suggest that the primes enhanced the auditory entrainment to the temporal structure of speech in the target sentence, which in turn enhanced its intelligibility, the current design does not allow us to distinguish among the following interpretations of the prime effect: (1) speech

processing was enhanced by the pure repetition of the speech envelope; (2) the prime induced entrainment which was itself enough to improve the processing of the target sentence; (3) or both repetition and entrainment are responsible for the target intelligibility enhancement. In our opinion, the third interpretation is the most likely. Furthermore, replicating these results with electrophysiological methods is necessary. Regarding the interpretations (2) and (3), and specifically in the case of our group of 5th graders, the improvement of pseudoword recall by the presence of a prime could be providing further support to the idea that entrainment to the slow amplitude envelope of speech is important for speech intelligibility (Peelle and Davis, 2012; Peelle et al., 2013). Our behavioral results are in line with the predictions of multi-temporal resolution models (Poeppel, 2003; Ghitza, 2011; Giraud and Poeppel, 2012): entrainment to the slow oscillations of speech (by means of priming) could be supported through oscillatory mechanisms at low frequency bands (delta and theta bands), and this could in turn signal “acoustic landmarks” in which to anchor fast phonemic processing (gamma band). Moreover, the benefit obtained from the prime in the 5th graders in the noisy listening environment corroborates the hypothesis that the prime helped the children to selectively attend the target sentence, which facilitated phonemic information processing. Our results are in line with the neurophysiological hypothesis suggesting that the brain relies on acoustic information at low frequencies to segregate a target signal from a background noise and create a clearer neural representation of this signal (Zion Golumbic et al., 2012; Ding and Simon, 2014). However, Grade 2 children benefited less clearly from the prime in both the quiet and the noisy environments. In noisy listening situations in particular, this absence of main prime effect could stem from the overall poor performance of Grade 2 children. Moreover, the Grade by Listening condition interaction suggests that pseudoword recall performance was especially degraded by the babble noise for children in Grade 2 as compared to children in Grade 5. According to studies showing that noise affects the timing of neural responses at high frequency bands such as those recorded

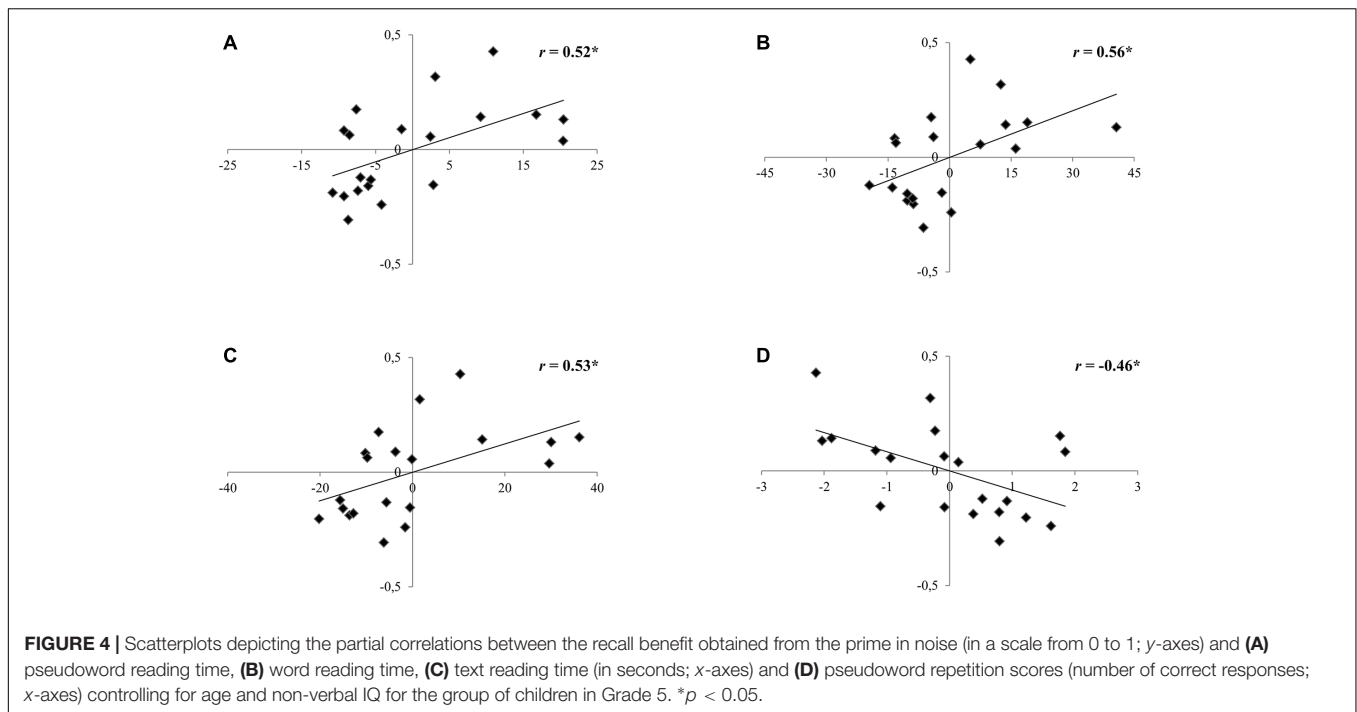
**TABLE 5 |** Mean accuracy (standard deviation in parenthesis) per Grade averaged across participants in the different positions and listening conditions of the Speech perception task.

	Grade 2 ( $n = 20$ )	Grade 5 ( $n = 20$ )
Quiet		
Position 1	0.67 (0.47)	0.93 (0.26)
Position 2	0.69 (0.46)	0.88 (0.33)
Position 3	0.87 (0.34)	0.93 (0.27)
Position 4	0.83 (0.38)	0.90 (0.30)
Noise		
Position 1	0.35 (0.48)	0.45 (0.50)
Position 2	0.23 (0.42)	0.40 (0.49)
Position 3	0.28 (0.45)	0.47 (0.50)
Position 4	0.23 (0.42)	0.39 (0.49)

**TABLE 6 |**  $R$ -values for two-tailed partial correlations between phonological and reading tasks and the benefit obtained from the prime in quiet and in noise controlling for non-verbal IQ and chronological age.

Benefit from the prime	Grade 2		Grade 5	
	In noise	In quiet	In noise	In quiet
Reading skills				
Word reading – Number of errors	–0.37	–0.40	–0.07	0.28
Word reading – Time (s)	–0.7	–0.17	0.56*	–0.11
Pseudoword reading – Number of errors	–0.15	–0.29	0.37	–0.29
Pseudoword reading – Time (s)	–0.30	–0.10	0.52*	–0.10
Text reading – Number of errors	0.01	0.19	0.09	–0.15
Text reading – Time (s)	–0.06	–0.23	0.53*	–0.04
Phonological skills				
Pseudoword repetition – Accuracy/12	–0.08	0.02	–0.46*	0.30
Phonemic deletion – Accuracy/12	–0.34	–0.32	0.06	0.15

PW, pseudoword. \* $p < 0.05$ .



at the brainstem (Anderson et al., 2010), severe temporal neural disruptions may have been associated with poor speech in noise perception in the younger age group. Therefore, the babble noise must have strongly prevented Grade 2 children to encode pseudoword phonemic content. These difficulties may reflect problems faced by young children to integrate speech information across various temporal scales (prosodic-syllabic and phonemic) in noisy listening conditions.

Another factor that could have contributed to group differences was the position of the target pseudoword in the sentence. A follow-up analysis revealed that while the performance of 5th graders was stable in the different positions across listening conditions, position had a strong effect on performance for the 2nd graders, especially with regard to positions 1 and 2 when compared to positions 3 and 4 (see Table 5 and Figure 3). On the one hand, this effect could be due to the fact that non-canonical orders of the sentence (pseudoword in positions 1 and 2) involved more complex language structures that younger children were not able to process fully automatically. On the other hand, group differences in the temporal deployment of attention might be at the base of this result. The 5th graders outperformed their younger peers when the target appeared in the initial positions of the sentence, supporting the idea that younger children might need more time to rapidly trigger auditory engagement to speech. An interesting follow-up of these results would be to further manipulate the time of appearance of the target pseudoword to explore the time course of attentional deployment over speech, which, at the neural level, could be reflecting the alternation between attentional engagement and disengagement moments and low- and high-excitability moments of neural tissue (Schroeder and Lakatos, 2009; Calderone et al., 2014; Lallier et al., 2017).

Lastly, it is important to note children in our speech perception task were instructed to repeat pseudowords. Although non-word repetition is a classical and extensively used task to measure phonological short term memory in children (e.g., Dollaghan and Campbell, 1998), it is a complex task involving auditory, phonological and speech-motor output processes (Gathercole, 2006). Therefore, one could ask whether the prime effects observed were the consequence of a boost all of these components rather than just speech intelligibility. Nevertheless, the main aim of the study was to test if a purely perceptual auditory prime would boost speech perception, and further studies are needed to disentangle the effect of auditory primes on phonological, auditory and motor output skills separately.

### Relation between Reading and Entrainment to Slow Speech Oscillations

Across grades, different correlation patterns were found between speech perception task performance, reading scores, and phonological scores. While no correlation was found within the 2nd graders group, phonological and reading measures correlated within the group of 5th graders. Moreover, a negative correlation was present between pseudoword repetition scores and the benefit obtained from the prime in the noisy condition, illustrating that children who were worse at repeating pseudowords benefited more from the prime that presumably facilitated the recall of the target pseudoword. Moreover, the slower readers among the 5th graders were those obtaining a greater benefit from the prime in noisy listening conditions. This result goes in line with recent theories of dyslexia (“temporal sampling framework,” Goswami, 2011) such that low reading abilities could be related to deficient sensitivity to slow speech

amplitude modulations. In particular, children with good reading skills received a smaller benefit from the prime, most likely because the repetition of the amplitude envelope was redundant. On the contrary, the mere repetition of the frequencies of speech below 8 Hz improved speech intelligibility for children with less sensitivity to amplitude changes, who also were the ones with lower reading skills.

Importantly, these correlations were found for speech perception in the noisy conditions only. This is in line with reports on significant links between speech perception in noise and reading (Wible et al., 2002; Bradlow et al., 2003; Ziegler et al., 2009; Boets et al., 2011), while such link between reading and speech perception has not been consistently found in quiet listening settings (Brady et al., 1983; Bradlow et al., 2003). In fact, since fast (phonemic) and slow (syllabic and prosodic) information is fully available in quiet listening conditions, a few processing resources may be sufficient to overcome and hide the auditory deficits of poorer readers, at least behaviourally, which might manifest if the processing load (i.e., added noise) to perform the task increases (Soroli et al., 2010).

## CONCLUSION

The present results showed that prior availability of the amplitude envelope of speech (i.e., via priming) at prosodic and syllabic rates (below 8 Hz) enhances speech intelligibility, and that this effect is modulated by chronological age. Although our behavioral results are in line with the predictions of multitemporal resolution models of speech, this relation remains speculative; more evidence is needed to corroborate the idea that oscillatory entrainment is causally related to behavior, since the possibility that this activity is a mere by-product of other underlying processes remains open. Importantly, we observed that, only in the case of the older children (when the gap

between children with different reading skills should start to be really meaningful), performance on the speech perception task significantly contributed to reading skills, such as the poorer the reading, the higher the prime benefit for processing speech in noisy listening conditions.

## ETHICS STATEMENT

This study was carried out in accordance with the recommendations of BCBL Ethics Committee with written informed consent from all subjects, parents, or tutors. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the BCBL Ethics Committee.

## AUTHOR CONTRIBUTIONS

PR-L, MaL, and MM wrote the manuscript. PR-L conducted the experiment and analyzed the data. PR-L and MiL prepared the stimuli.

## FUNDING

This work was supported by the Severo Ochoa program grant [SEV-2015-049]; MaL was supported by the Spanish Ministry of Economy and Competitiveness [PSI2015-65338-P].

## ACKNOWLEDGMENTS

The authors wish to thank Doug Davidson for his insightful suggestions on data analysis, and the school personnel, the children and their parents for participating in this study.

## REFERENCES

- Abrams, D. A., Nicol, T., Zecker, S., and Kraus, N. (2008). Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *J. Neurosci.* 28, 3958–3965. doi: 10.1523/JNEUROSCI.0187-08.2008
- Abrams, D. A., Nicol, T., Zecker, S., and Kraus, N. (2009). Abnormal cortical processing of the syllable rate of speech in poor readers. *J. Neurosci.* 29, 7686–7693. doi: 10.1523/JNEUROSCI.5242-08.2009
- Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., and Merzenich, M. M. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 98, 13367–13372. doi: 10.1073/pnas.201400998
- Aiken, S. J., and Picton, T. W. (2008). Human cortical responses to the speech envelope. *Ear. Hear.* 29, 139–157. doi: 10.1097/AUD.0b013e31816453dc
- Anderson, S., Skoe, E., Chandrasekaran, B., and Kraus, N. (2010). Neural timing is linked to speech perception in noise. *J. Neurosci.* 30, 4922–4926. doi: 10.1523/JNEUROSCI.0107-10.2010
- Baayen, H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modelling with crossed random effects for subjects and items. *J. Mem. Lang.* 59, 390–412. doi: 10.1016/j.jml.2007.12.005
- Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang.* 68, 255–278. doi: 10.1016/j.jml.2012.11.001
- Bates, D., Maechler, M., Bolker, B., and Walker, S. (2013). *Linear Mixed-Effect Models Using Eigen and S4. R Package Version: 1.0-5*. Available at: <http://cran.rproject.org/package=lme4>
- Boemio, A., Fromm, S., Braun, A., and Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat. Neurosci.* 8, 389–395. doi: 10.1038/nn1409
- Boets, B., Vandermosten, M., Poelmans, H., Luts, H., Wouters, J., and Ghesquière, P. (2011). Preschool impairments in auditory processing and speech perception uniquely predict future reading problems. *Res. Dev. Disabil.* 32, 560–570. doi: 10.1016/j.ridd.2010.12.020
- Bourguignon, M., De Tieghe, X., de Beeck, M. O., Ligot, N., Paquier, P., Van Bogaert, P., et al. (2013). The pace of prosodic phrasing couples the listener's cortex to the reader's voice. *Hum. Brain Mapp.* 34, 314–326. doi: 10.1002/hbm.21442
- Bradlow, A. R., Kraus, N., and Hayes, E. (2003). Speaking clearly for children with learning disabilities: sentence perception in noise. *J. Speech Lang. Hear. Res.* 46, 80–97. doi: 10.1044/1092-4388(2003)007
- Brady, S., Shankweiler, D., and Mann, V. (1983). Speech perception and memory coding in relation to reading ability. *J. Exp. Child Psychol.* 35, 345–367. doi: 10.1016/0022-0965(83)90087-5
- Brungart, D. S., Simpson, B. D., Ericson, M. A., and Scott, K. R. (2001). Informational and energetic masking effects in the perception of multiple

- simultaneous talkers. *J. Acoust. Soc. Am.* 110, 2527–2538. doi: 10.1121/1.1408946
- Calderone, D. J., Lakatos, P., Butler, P. D., and Castellanos, F. X. (2014). Entrainment of neural oscillations as a modifiable substrate of attention. *Trends Cogn. Sci.* 18, 300–309. doi: 10.1016/j.tics.2014.02.005
- Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., et al. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. *Science* 313, 1626–1628. doi: 10.1126/science.1128115
- Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavský, M., Onochie-Quintanilla, E., et al. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychol. Sci.* 23, 678–686. doi: 10.1177/0956797611434536
- Cason, N., and Schön, D. (2012). Rhythmic priming enhances the phonological processing of speech. *Neuropsychologia* 50, 2652–2658. doi: 10.1016/j.neuropsychologia.2012.07.018
- Chait, M., Greenberg, S., Arai, T., Simon, J. Z., and Poeppel, D. (2015). Multi-time resolution analysis of speech: evidence from psychophysics. *Front. Neurosci.* 9:214. doi: 10.3389/fnins.2015.00214
- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *J. Acoust. Soc. Am.* 25, 975–979. doi: 10.1121/1.1907229
- Corriveau, K., Pasquini, E., and Goswami, U. (2007). Basic auditory processing skills and specific language impairment: a new look at an old hypothesis. *J. Speech Lang. Hear. Res.* 50, 647–666. doi: 10.1044/1092-4388(2007/046)
- Cuetos, F., Rodríguez, B., Ruano, E., and Arribas, D. (2007). *PROLEC-R: Batería de Evaluación de Los Procesos Lectores Revisada*. Madrid: TEA.
- Ding, N., and Simon, J. Z. (2012). Emergence of neural encoding of auditory objects while listening to competing speakers. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11854–11859. doi: 10.1073/pnas.1205381109
- Ding, N., and Simon, J. Z. (2014). Cortical entrainment to continuous speech: functional roles and interpretations. *Front. Hum. Neurosci.* 8:311. doi: 10.3389/fnhum.2014.00311
- 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. doi: 10.1016/j.neuropsychologia.2012.03.007
- Dollaghan, C., and Campbell, T. F. (1998). Nonword repetition and child language impairment. *J. Speech Lang. Hear. Res.* 41, 1136–1146. doi: 10.1044/jslhr.4105.1136
- Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., and Carreiras, M. (2013). EsPal: one-stop shopping for Spanish word properties. *Behav. Res. Methods* 45, 1246–1258. doi: 10.3758/s13428-013-0326-1
- Edwards, J., Beckman, M. E., and Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *J. Speech Lang. Hear. Res.* 47, 421–436. doi: 10.1044/1092-4388(2004/034)
- Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Appl. Psycholinguist* 27, 513–543. doi: 10.1017/S0142716406060383
- Ghitza, O. (2011). Linking speech perception and neurophysiology: speech decoding guided by cascaded oscillators locked to the input rhythm. *Front. Psychol.* 2:130. doi: 10.3389/fpsyg.2011.00130
- Ghitza, O. (2012). On the role of theta-driven syllabic parsing in decoding speech: intelligibility of speech with a manipulated modulation spectrum. *Front. Psychol.* 3:238. doi: 10.3389/fpsyg.2012.00238
- Ghitza, O., and Greenberg, S. (2009). On the possible role of brain rhythms in speech perception: intelligibility of time-compressed speech with periodic and aperiodic insertions of silence. *Phonetica* 66, 113–126. doi: 10.1159/000208934
- Giraud, A. L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S., and Laufs, H. (2007). Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron* 56, 1127–1134. doi: 10.1016/j.neuron.2007.09.038
- Giraud, A. L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nat. Neurosci.* 15, 511–517. doi: 10.1038/nn.3063
- Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends Cogn. Sci.* 15, 3–10. doi: 10.1016/j.tics.2010.10.001
- Goswami, U., Cumming, R., Chait, M., Huss, M., Mead, N., Wilson, A. M., et al. (2016). Perception of filtered speech by children with developmental dyslexia and children with specific language impairments. *Front. Psychol.* 7:791. doi: 10.3389/fpsyg.2016.00791
- Goswami, U., Gerson, D., and Astruc, L. (2010). Amplitude envelope perception, phonology and prosodic sensitivity in children with developmental dyslexia. *Read. Writ.* 23, 995–1019. doi: 10.1007/s11145-009-9186-6
- Goswami, U., and Leong, V. (2013). Speech rhythm and temporal structure: converging perspectives. *Lab. Phonol.* 4, 67–92. doi: 10.1515/lp-2013-0004
- Goswami, U., Power, A. J., Lallier, M., and Facoetti, A. (2014). Oscillatory “temporal sampling” and developmental dyslexia: toward an over-arching theoretical framework. *Front. Hum. Neurosci.* 8:904. doi: 10.3389/fnhum.2014.00904
- Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S., et al. (2002). Amplitude envelope onsets and developmental dyslexia: a new hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 99, 10911–10916. doi: 10.1073/pnas.122368599
- Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., et al. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biol.* 11:e1001752. doi: 10.1371/journal.pbio.1001752
- Hämäläinen, J., Leppänen, P. H. T., Torppa, M., Müller, K., and Lyytinen, H. (2005). Detection of sound rise time by adults with dyslexia. *Brain Lang.* 94, 32–42. doi: 10.1016/j.bandl.2004.11.005
- Hämäläinen, J. A., Rupp, A., Soltész, F., Szücs, D., and Goswami, U. (2012). Reduced phase locking to slow amplitude modulation in adults with dyslexia: an MEG study. *Neuroimage* 59, 2952–2961. doi: 10.1016/j.neuroimage.2011.09.075
- Hazan, V., Messaoud-Galusi, S., Rosen, S., Nouwens, S., and Shakespeare, B. (2009). Speech perception abilities of adults with dyslexia: is there any evidence for a true deficit? *J. Speech Lang. Hear. Res.* 52, 1510–1529. doi: 10.1044/1092-4388(2009/08-0220)
- Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402. doi: 10.1038/nrn2113
- Houtgast, T., and Steeneken, H. J. (1985). A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. *J. Acoust. Soc. Am.* 77, 1069–1077. doi: 10.1121/1.392224
- Hulme, C., Snowling, M., Caravolas, M., and Carroll, J. (2005). Phonological skills are (probably) one cause of success in learning to read: a comment on Castles and Coltheart. *Sci. Stud. Read.* 9, 351–365. doi: 10.1207/s1532799xssr0904\_2
- Kerlin, J. R., Shahin, A. J., and Miller, L. M. (2010). Attentional gain control of ongoing cortical speech representations in a “cocktail party”. *J. Neurosci.* 30, 620–628. doi: 10.1523/JNEUROSCI.3631-09.2010
- Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., and Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320, 110–113. doi: 10.1126/science.1154735
- Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., and Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *J. Neurophysiol.* 94, 1904–1911. doi: 10.1152/jn.00263.2005
- Lallier, M., Donnadieu, S., and Valdois, S. (2012). Investigating the role of visual and auditory search in reading and developmental dyslexia. *Front. Hum. Neurosci.* 7:597. doi: 10.3389/fnhum.2013.00597
- Lallier, M., Molinaro, N., Lizarazu, M., Bourguignon, M., and Carreiras, M. (2017). Amodal atypical neural oscillatory activity in dyslexia: a cross-linguistic perspective. *Clin. Psychol. Sci.* 5, 379–401. doi: 10.1177/2167702616670119
- Lallier, M., Valdois, S., Lassus-Sangosse, D., Prado, C., and Kandel, S. (2014). Impact of orthographic transparency on typical and atypical reading development: evidence in French-Spanish bilingual children. *Res. Dev. Disabil.* 35, 1177–1190. doi: 10.1016/j.ridd.2014.01.021
- Lehongre, K., Morillon, B., Giraud, A. L., and Ramus, F. (2013). Impaired auditory sampling in dyslexia: further evidence from combined fMRI and EEG. *Front. Hum. Neurosci.* 7:454. doi: 10.3389/fnhum.2013.00454
- Leong, V., and Goswami, U. (2015). Acoustic-emergent phonology in the amplitude envelope of child-directed speech. *PLoS ONE* 10:e0144411. doi: 10.1371/journal.pone.0144411
- Luo, H., and Poeppel, D. (2012). Cortical oscillations in auditory perception and speech: evidence for two temporal windows in human auditory cortex. *Front. Psychol.* 3:170. doi: 10.3389/fpsyg.2012.00170



- Molinari, N., Lizarazu, M., Lallier, M., Bourguignon, M., and Carreiras, M. (2016). Out-of-synchrony speech entrainment in developmental dyslexia. *Hum. Brain Mapp.* 37, 2767–2783. doi: 10.1002/hbm.23206
- Nourski, K. V., Reale, R. A., Oya, H., Kawasaki, H., Kovach, C. K., Chen, H., et al. (2009). Temporal envelope of time-compressed speech represented in the human auditory cortex. *J. Neurosci.* 29, 15564–15574. doi: 10.1523/JNEUROSCI.3065-09.2009
- Obleser, J., Eisner, F., and Kotz, S. A. (2008). Bilateral speech comprehension reflects differential sensitivity to spectral and temporal features. *J. Neurosci.* 28, 8116–8123. doi: 10.1523/JNEUROSCI.1290-08.2008
- Obleser, J., Herrmann, B., and Henry, M. J. (2012). Neural oscillations in speech: don't be enslaved by the envelope. *Front. Hum. Neurosci.* 6:250. doi: 10.3389/fnhum.2012.00250
- Peelle, J. E., and Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Front. Psychol.* 3:320. doi: 10.3389/fpsyg.2012.00320
- Peelle, J. E., Gross, J., and Davis, M. H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cereb. Cortex* 23, 1378–1387. doi: 10.1093/cercor/bhs118
- Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech Commun.* 41, 245–255. doi: 10.1016/S0167-6393(02)00107-3
- Poeppel, D. (2014). The neuroanatomic and neurophysiological infrastructure for speech and language. *Curr. Opin. Neurobiol.* 28, 142–149. doi: 10.1016/j.conb.2014.07.005
- Power, A. J., Mead, N., Barnes, L., and Goswami, U. (2013). Neural entrainment to rhythmic speech in children with developmental dyslexia. *Front. Hum. Neurosci.* 7:777. doi: 10.3389/fnhum.2013.00777
- Przybylski, L., Bedoin, N., Krifi-Papoz, S., Herbillon, V., Roch, D., Léculier, L., et al. (2013). Rhythmic auditory stimulation influences syntactic processing in children with developmental language disorders. *Neuropsychology* 27, 121–131. doi: 10.1037/a0031277
- R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 336, 367–373. doi: 10.1098/rstb.1992.0070
- Rothermich, K., Schmidt-Kassow, M., and Kotz, S. A. (2012). Rhythm's gonna get you: regular meter facilitates semantic sentence processing. *Neuropsychologia* 50, 232–244. doi: 10.1016/j.neuropsychologia.2011.10.025
- Saoud, H., Josse, G., Bertasi, E., Truy, E., Chait, M., and Giraud, A. L. (2012). Brain-speech alignment enhances auditory cortical responses and speech perception. *J. Neurosci.* 32, 275–281. doi: 10.1523/JNEUROSCI.3970-11.2012
- Schroeder, C. E., and Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci.* 32, 9–18. doi: 10.1016/j.tins.2008.09.012
- Soroli, E., Szenkovits, G., and Ramus, F. (2010). Exploring dyslexics' phonological deficit III: foreign speech perception and production. *Dyslexia* 16, 318–340. doi: 10.1002/dys.415
- Stanovich, K. E., Cunningham, A. E., and Cramer, B. B. (1984). Assessing phonological awareness in kindergarten children: issues of task comparability. *J. Exp. Child Psychol.* 38, 175–190. doi: 10.1016/0022-0965(84)90120-6
- Suppes, P., Han, B., and Lu, Z. L. (1998). Brain-wave recognition of sentences. *Proc. Natl. Acad. Sci. U.S.A.* 95, 15861–15866. doi: 10.1073/pnas.95.26.15861
- Suppes, P., Lu, Z. L., and Han, B. (1997). Brain wave recognition of words. *Proc. Natl. Acad. Sci. U.S.A.* 94, 14965–14969. doi: 10.1073/pnas.94.26.14965
- Wible, B., Nicol, T., and Kraus, N. (2002). Abnormal neural encoding of repeated speech stimuli in noise in children with learning problems. *Clin. Neurophysiol.* 113, 485–494. doi: 10.1016/S1388-2457(02)00017-2
- Yarkoni, T., Balota, D., and Yap, M. (2008). Moving beyond Coltheart's N: a new measure of orthographic similarity. *Psychon. Bull. Rev.* 15, 971–979. doi: 10.3758/PBR.15.5.971
- Zatorre, R. J., and Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11, 946–953. doi: 10.1093/cercor/11.10.946
- Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faisca, L., et al. (2010). Orthographic depth and its impact on universal predictors of reading: a cross-language investigation. *Psychol. Sci.* 21, 551–559. doi: 10.1177/0956797610363406
- Ziegler, J. C., Pech-Georgel, C., George, F., and Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Dev. Sci.* 12, 732–745. doi: 10.1111/j.1467-7687.2009.00817.x
- Zion Golumbic, E. M., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., et al. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". *Neuron* 77, 980–991. doi: 10.1016/j.neuron.2012.12.037
- Zion Golumbic, E. M., Poeppel, D., and Schroeder, C. E. (2012). Temporal context in speech processing and attentional stream selection: a behavioral and neural perspective. *Brain Lang.* 122, 151–161. doi: 10.1016/j.bandl.2011.12.010

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Ríos-López, Molnar, Lizarazu and Lallier. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.