CATÓLICA
INSTITUTO DE CIÊNCIAS DA SAÚDE
LISBOA · PORTO · VISEU

THE EFFECT OF LITERACY IN THE SPEECH TEMPORAL MODULATION STRUCTURE

Dissertação apresentada à Universidade Católica Portuguesa para obtenção do grau de mestre em

NEUROPSICOLOGIA

Por

João Pedro Fonseca de Araújo

Lisboa – 2016

CATÓLICA
INSTITUTO DE CIÊNCIAS DA SAÚDE
LISBOA·PORTO·VISEU

THE EFFECT OF LITERACY IN THE SPEECH TEMPORAL MODULATION STRUCTURE

Dissertação apresentada à Universidade Católica Portuguesa para obtenção do grau de mestre em

NEUROPSICOLOGIA

Por

João Pedro Fonseca de Araújo

Sob a orientação do Professor Doutor Alexandre Castro-Caldas

Lisboa – 2016

Abstract

The temporal modulation structure of adult-directed speech is conceptualised as a modulation hierarchy comprising four temporal bands, delta, 1 – 3 Hz, theta, 4 – 8 Hz, beta, 15 – 30 Hz, and low gamma, 30 – 50 Hz. Neuronal oscillatory entrainment to amplitude modulations (AMs) in these four bands may provide a basis for speech encoding and parsing the continuous signal into linguistic units (delta – syllable stress patterns, theta – syllables, beta – onset-rime units, low gamma – phonetic information). While adult-directed speech is theta-dominant and shows tighter theta-beta/low gamma phase alignment, infant-directed speech is delta-dominant and shows tighter delta-theta phase alignment. Although this change in the speech representations could be maturational, it was hypothesized that literacy may also influence the structure of speech. In fact, literacy and schooling are known to change auditory speech entrainment, enhancing phonemic specification and augmenting the phonological detail of the lexicon's representations. Thus, we hypothesized that a corresponding difference in speech production could also emerge. In this work, spontaneous speech samples were recorded from literate (with lower and higher literacy) and illiterate subjects and their energy modulation spectrum across delta, theta and beta/low gamma AMs as well as the phase synchronization between nested AMs analysed. Measures of the participants' phonology skills and vocabulary were also retrieved and a specific task to confirm the sensitivity to speech rhythm of the analysis method used (S-AMPH) was conducted. Results showed no differences in the energy of delta, theta and beta/low gamma AMs in spontaneous speech. However, phase alignment between slower and faster speech AMs was significantly enhanced by literacy, showing moderately strong correlations with the phonology measures and literacy. Our data suggest that literacy affects not only cortical entrainment and speech perception but also the physical/rhythmic properties of speech production.

Keywords: Literacy, Speech Envelope, Modulation Spectrum, AM Synchronisation, Psycholinguistics

Resumo

A modulação temporal do discurso dirigido a adultos é conceptualizado como uma hierarquia de modulações em quatro bandas temporais: delta, 1 – 3 Hz, theta, 4 – 8 Hz, beta, 15 – 30 Hz, e low gamma, 30 – 50 Hz. A sincronização das oscilações neuronais nestas quatro bandas pode providenciar a base para a codificação e análise de um sinal contínuo em unidades linguísticas (delta – força silábica, theta – sílabas, beta – arranque/rima, low gamma – informação fonética). Enquanto o discurso dirigido a adultos é de um ritmo predominantemente theta e mostra um forte alinhamento entre bandas theta e beta/low gamma, discurso dirigido a crianças é predominantemente de um ritmo delta e mostra maiores sincronizações entre bandas delta e theta. Apesar das diferenças nas representações do discurso poderem resultar de processos maturacionais, foi hipotetizado que a literacia também poderia influenciar as características rítmicas do discurso. De facto, a literacia afecta o processamento auditivo da linguagem, além de desenvolver a consciência fonémica e aumentar o detalhe fonológico das representações lexicais. Neste estudo foram gravadas amostras de discurso espontâneo de sujeitos letrados (alta e baixa escolarização) e iletrados. Os espectros de modulação de energia nas bandas de interesse foram analisados bem como a sincronização das bandas delta-theta e theta-beta/ low gamma. Foram recolhidas medidas de consciência fonológica e vocabulário e foi realizada também uma tarefa para confirmar a sensibilidade do modelo de análise (S-AMPH) ao ritmo do discurso. A análise revelou ausência de diferenças na energia nas modulações delta, theta ou beta/low gamma no discurso espontâneo. Contudo, a sincronização entre as bandas aumentou significativamente com a literacia, revelando uma correlação moderada com as medidas de fonologia, vocabulário e literacia. Sendo assim, a literacia afecta não só a sincronização cortical e à linguagem falada mas também as propriedades físicas e rítmicas da produção do discurso.

Palavras-Chave: Literacia, Envelope do Discurso, Espectro de Modulação, Sincronização de AMs, Psicolinguística

Agradecimentos/Aknowledgements

Em baixo segue a minha lista de agradecimentos às pessoas que tornaram este trabalho não só possível mas também uma aventura na qual aprendi tanto:

Ao professor Alexandre Castro-Caldas, cuja orientação e ajuda indispensáveis me fizeram querer chegar mais além e aguçaram ainda mais o meu gosto pela investigação. Obrigado por acreditar em mim, por vezes mais do que eu próprio, e por me conceder a oportunidade de trabalhar nesta área inovadora da neurociência da linguagem.

To the people that welcomed me at the CNE when I visited to learn about the S-AMPH model, including Mrs Nichola Daily who helped me finding a place to stay. I would like to thank Dr Sheila Flanagan for teaching me most of what I know about the model, dealing with all my questions while I was in Cambridge, for introducing me to the theoretical aspects of psychoacoustics and for the valuable help on the data analysis. Also, this thesis would definitely not have been possible without the constant guidance of Professor Usha Goswami. I feel blessed for having had your support and for the opportunity of learning from you since the day we met.

Às entidades que aceitaram colaborar neste estudo, nomeadamente a Universidade Sénior da Junta de Freguesia de Alcântara (especialmente à Dra Joana Troni e à professora Amélia Simões Figueiredo), o Centro de Dia de Aveiras, à Associação de Reformados, Pensionistas e Idosos “Os Bispinhos”, ao Centro Social e Paroquial São Tomás de Aquino e aos participantes de norte a sul do país que deste estudo fizeram parte. Um agradecimento especial às “avós” Maria e Céu e aos “avôs” Manel e Albertino por me apresentarem a muitos dos sujeitos que participaram neste trabalho.

À minha família pelo seu incansável apoio, aos meus pais e à tia Dete que me deram algumas boleias indispensáveis e cujos contactos foram imprescindíveis. Um agradecimento especial à minha namorada Raquel Cunha, a melhor do mundo, pelo carinho e por ser um pilar fundamental para mim e me motivar a ser melhor todos os dias. Aos meus amigos de sempre, aos que fiz no mestrado e na AIESEC quero também agradecer pelos momentos de maior descontração tão necessários no desenvolvimento saudável de uma tese.

Aos que me ajudaram no processo de criação do meu paradigma, nomeadamente à mestre Ana Filipa Miranda pelo “know-how” e ajuda na preparação do método de entrevista, às professoras Filipa Ribeiro e Maria Emília Santos pela ajuda na selecção dos instrumentos de avaliação e ao Miguel Morgado por me ajudar a compreender algumas questões teóricas da física do discurso.

À Universidade Católica de Lisboa e a todos os que, de alguma forma, contribuiram para este trabalho, muito obrigado!

Contents

List of Tables

List of Figures

Introduction

In previous literature, the neural basis of speech processing has been approached not only by neuroanatomical models (like the model of Friederici, 2002) but more recently by the study of oscillatory activity predominantly on the auditory cortex (Giraud & Poeppel, 2012). Indeed, the idea that the auditory cortex could support speech perception and production dates back to the 1870's and it was built on the observation of typical symptoms revealed by subjects with lesions in the left superior temporal gyrus (Wernicke, 1874/1977). In subsequent studies it was reported that lesions in that area were more related to speech production deficits, although functional imaging studies would still highlight superior temporal regions when speech stimuli was presented (Hickok & Poeppel, 2007 for a review). In their work, Hickok and Poeppel suggested a mapping of the cortical fields involved in language comprehension and production independent of specific language tasks (Hickok & Poeppel, 2000) and proposed a dual-stream neuroanatomical model for speech processing (Hickok & Poeppel, 2004). Further, they have extended that model, with a particular focus on the bilateral parallel computations and the multi-time resolution processing of the ventral stream (Hickok & Poeppel, 2007). Finally, Giraud and Poeppel (2012) published their new oscillation-based functional model, focusing on the processing of speech at different time intervals (corresponding to different linguistic units) and on the role of oscillations in segregating sensory information and organizing neural spike timing. In this model, the right auditory cortex is of crucial importance, particularly for the integration of information in longer temporal windows (Poeppel, 2003; Poeppel, 2014). On the other hand, the left auditory areas have shown to be important for the encoding of linguistic information in short temporal windows according to studies of speech perception and production (Hickok & Poeppel, 2000).

Recent studies with dyslexic subjects and children with specific language impairment have been addressing how literacy and phonology variables may impact this representation of speech temporal features (Molinaro et al., 2016; Power et al., 2013; Goswami et al., 2002; Lehongre et al., 2011; Goswami et al., 2016). However, no study so far has addressed the effect of literacy in the speech temporal structure by testing groups of adults with different levels of schooling – including no schooling at all (illiterates) –which is the research focus of this work. In Portugal, 50 years ago, it was common for the older daughters (and, in some cases, their male siblings) of a family to be engaged in household or work-related activities since a very young age and, therefore, they did not enter school. Because of that previous social/cultural background, there are still some illiterate older adults in Portugal (particularly in rural areas) and the study of this population has been enlightening to understand how literacy affects phonology and the brain (Castro-Caldas et al., 1998, 1999; Castro-Caldas & Reis, 2000; Petersson et al., 1999, 2000). Furthermore, a better insight on how literacy changes the speech representations may also lead to a better understanding of aphasia and how similar brain damage may result in clinical distinct manifestations of language impairment in literates and illiterates (Matute, 1986).

Literature Review

Speech signal processing: when psychoacoustics meets neuroscience

Temporal rhythm is crucial for the neural processing when encoding or transmitting sensory information (Goswami, 2016). In fact, the speech signal is rich in temporal and frequency patterns and can be operationalized as the summation of several of these frequency bands fluctuating in intensity over time (Elliot & Theunissen, 2009; Goswami, 2011). Speech sound involves only the lower part of audible frequencies (20 Hz – 8 kHz) where different information is carried by different timescales: faster modulations between ~70Hz – 1kHz define pitch and inter-aural perception / sound localization – usually called temporal fine structure – and slower modulations between ~1Hz – 10Hz are heard as a sequence of discrete events that include syllables and words – usually called temporal envelope (Arnal, Poeppel & Giraud, 2015). The ability to utilize the cues of the envelope for speech comprehension was already suggested by Shannon et al. (1995) that reported slowly varying temporal information (< 16 Hz and < 50 Hz) could lead to relatively high recognition performance under conditions of reduced spectral information. Besides, it is reported excellent speech comprehension when the signal is divided into four frequency bands (16, 50, 160 and 500 Hz) using its temporal envelopes to modulate noise of the same bandwidth. Therefore, the temporal envelope yields sufficient information for speech comprehension and peaks at approximately 4 – 6 Hz, the corresponding frequency band of the syllabic structure of adult-directed speech (Giraud & Poeppel, 2012; Greenberg et al., 2003). Moreover, the envelope includes the temporal modulations matching linguistic units such as prosody changes (1 – 4 Hz), syllables (4 – 8 Hz), onset-rhymes (15 – 30 Hz) and phonemes (30 – 45 Hz) (Peterson & Lehiste, 1960; Greenberg, 2006; Ghitza & Greenberg, 2009; Poeppel, 2014; Goswami, 2016). Chait et al. (2015) report that when the envelope is presented with its faster and slower modulations the improvement of comprehension is greater than the simple linear combination of the performance when only slower or rapid modulations are presented. This may suggest that a binding process of these linguistic units of multiple timescales occurs creating a representation of the sensory input that is more than the sum of its parts (Chait et al., 2015).

Interestingly, the frequency of neural oscillations (synchronous activity of neuronal ensembles) when the brain is at rest matches the frequency of the previously referred linguistic features (Arnal, Poeppel & Giraud, 2015). Neural oscillations control neuronal excitability and are involved in mechanisms of selective attention (particularly the gamma activity), temporally constraining (either towards a “rhythmic” or a “continuous” mode of operation) the sampling of the sensory input (Schroeder & Lakatos, 2008). According to multi-time resolution models of speech processing (see Giraud & Poeppel, 2012), oscillations in the auditory cortex interact with the neuronal (spiking) activity generated by the speech input, discretizing the signal in at least two timescales linked to different activity bands: *theta* (4 Hz – 8Hz) and *gamma* (30 Hz – 80 Hz) associated with the processing of syllables and phonemes respectively. The oscillatory activity on a third band – *delta* (1 Hz – 3 Hz) – is also related with speech intelligibility and it was hypothesized that this band tracks the acoustic rhythm and prosodic information (Ding & Simon, 2014; Goswami, 2011; Goswami, 2015; Ghitza & Greenberg, 2009). In a first stage of processing, after the encoding of the spectral-temporal properties of the speech signal, the input saliences cause a phase resetting of the oscillatory activity and while delta and theta bands entrain (phase lock) to the stimulus envelope, gamma regulates the spike patterns of neuronal activity (Poeppel, 2014; Giraud & Poeppel, 2012). These different nested oscillations in the auditory cortex display a hierarchical organization so that delta phase modulates theta amplitude and theta phase modulates gamma phase and amplitude (Lakatos et al., 2005; Canolty et al., 2006).

In short, the cortical oscillations entrain their activity at corresponding timescales in the signal to different modulations that regulate each other hierarchically. Then, oscillations regulate the neuronal excitability aligning with the speech signal temporal properties binding the information to create the final speech percept. But although this cortical oscillations show a synergistic relationship, the temporal processing at different timescales is associated with asymmetric speech sampling with stronger activations in the left than in the right hemisphere (Giraud & Poeppel, 2012). Moreover, the left auditory cortex has a higher proportion of cell groups that oscillate at higher (gamma) frequencies and the right auditory cortex has a higher proportion of cell groups that oscillate at lower (delta and theta) frequencies (Arnal, Poeppel & Giraud, 2015). Functional models of speech processing postulate that gamma sampling networks may be localized in both hemispheres while theta range networks may be strongly represented in the right hemisphere (Hickok & Poeppel, 2007 for a review). In fact, this sampling asymmetry has been reported in the superior temporal sulcus, being hypothesized that left and right superior temporal sulcus would receive input of different timescales (from the superior temporal gyrus) through intra-hemispheric and transcallosal pathways (Boemio et al., 2005).

From delta to theta processing: Maturation vs Literacy

Oscillatory activity in the brain seems to develop in early maturational processes, being already present in utero (Sieben et al., 2013 for a review). The emergence of faster hierarchically organized oscillations (essential for syllable and phoneme processing) might even be related with the early development of superficial (II and III) cortical layers (Gireesh & Plenz, 2008). It was also reported that neonates show sensitivity to supra-segmental information in speech and that by 6 months of age infants show adult-like frequency selectivity extracting information about spectral shape and temporal resolution of the speech signal (Bailey & Snowling, 2002). Slow speech AMs like prosody/intonation changes and rhythmic cues seem to be essential for young infants in tasks of language discrimination (Mehler et al., 1988; Ramus, 2002). In fact, higher energy in the delta band as well as a higher phase alignment between delta and theta bands may have a key role in language acquisition as both are present in motherese (Leong et al., 2014). Children with dyslexia, a developmental disorder characterized by difficulties in literacy acquisition, also show impaired entrainment and atypical right hemisphere response to delta modulations (Power et al., 2013; Power et al., 2016). Though, it was reported that in subjects with dyslexia the development from childhood to adulthood does not affect this reduced speech-brain synchronization (Molinaro et al., 2016).

However, typical adult-directed speech peaks at a theta (syllable) timescale and even though faster AMs like gamma (phonemes) are important for phonological encoding (even for speakers of languages that do not orthographically code this units like Mandarin, see Qu, Damian & Kazanina, 2012), it is the theta phase pattern that is correlated with speech intelligibility in adults (Luo & Poeppel, 2007). Interestingly, studies of oscillatory entrainment to speech by typically-developing children show a relationship between theta entrainment and reading skill (Power et al., 2012). Moreover, the performance in phonemic awareness tasks (deeply related to literacy, see next section) is strongly correlated with the auditory representation of syllables in dyslexic and normally-developing children (Power et al., 2013). Although dyslexics show no difference in theta entrainment strength when compared to controls, they do reveal a different (earlier) phase of entrainment in this timescale (Leong & Goswami, 2013). Summing up, on the one hand, speech processing seems to change from a delta to theta-based parsing over time from infancy to adulthood, what could suggest the influence of maturational processes. On the other hand, literacy and phonology measures seem to influence theta entrainment. In fact, the relationship of sensitivity to rhythmic / acoustic cues and literacy has already been studied. A correlation between reading acquisition and sensitivity to rhythmic properties such as slow AM rise times has been reported in child populations (Goswami et al., 2002). Moreover, a recent study has also suggested that performance in rhyme awareness tasks is related to reading level and sensitivity to slow (delta) amplitude modulation (AM) patterns (Leong & Goswami, 2016).

The relationship between literacy and phonology

One of the consequences of literacy acquisition is the training of phonological awareness – which can be broadly defined as the degree of sensitivity to the sound structure of oral language (Anthony & Francis, 2015). Moreover, one of the hallmarks of learning how to read is the acquisition of phoneme awareness – the ability to recognize, discriminate and manipulate the smallest linguistic units (phonemes) that constitute words (Anthony & Francis, 2015). In fact, even though pre-schoolers reveal good performance in phonological awareness tasks for syllables and onset-rhymes, awareness to phonemes is strictly dependent on the acquisition of literacy (Bowey & Francis, 1991). Previous studies with adults also show that illiterates are unable to manipulate (add or delete) phonemes in words and that even their performance in rhyme and syllable awareness is worse than ex-illiterates (Morais et al., 1979, 1986). A more recent study showed that performance in letter fluency and phoneme deletion tasks could discriminate between illiterates and semi-literates and showed the strongest correlations with letter knowledge in illiterates and word-reading in semi-literates (Loureiro et al., 2004). Moreover, in phonological fluency tasks, when illiterates are asked to say words starting with a syllable instead of a phoneme they improve their performance while literates show the opposite results (Kim et al., 2015).

Not only literacy is a determinant of phoneme awareness but it also seems to modulate speech processing and the access to speech representations. Castro-Caldas et al. (1998) reported that illiterates were worse than literates on non-word repetition tasks. Moreover, a subsequent study with greek subjects showed that was reading acquisition and not schooling that was related to those differences (Kosmidis et al., 2006). In these studies, it was hypothesized that literacy would be necessary for the full development of a “phonological route” of speech comprehension – a route that would bypass the auditory input lexicon and the speech output lexicon to perform a direct analysis at the phoneme level (see Ellis & Young, 1988; p.143-145 for an overview). In fact, lexicalization (transformation of a pseudoword into a meaningful word) accounted for more than 25% of the errors in this task in illiterates (Reis & Castro-Caldas, 1997). Literacy also changes the mental lexico-semantic organization. A couple of studies (Reis & Castro-Caldas, 1997; Petersson et al., 1999), using a task of word-pair memorization, showed that illiterates perform generally worse than their literate controls. Moreover, in this task, the biggest performance difference was on the phonologically-related pairs (e.g. *selo-pelo*; stamp-hair in English) that only differed in one phoneme and shared the same rhyme, suggesting that illiterates do not use phonological association between words as effectively as literates (Reis & Castro-Caldas, 1997; Petersson et al., 1999).

On the other hand, reading presents itself as a new modality of vocabulary acquisition, what can also affect phonology. According to Ziegler & Goswami (2005), in the early development of language, the grouping of lexical forms by phonological similarity is not refined. Indeed, the lexical forms may be grouped by onset-rhyme similarity (de Cara & Goswami, 2003). However, with the growth of vocabulary, lexical representations become integrated with more fine-grained phonological information, in a process that requires the learning of the discrete representations and the generalization of structural regularities present in the lexicon. This process is called lexical restructuring and results on the increase of phonological detail of word representations across both large and small linguistic units (Ziegler & Goswami, 2005). According to this view, dense phonologic neighbourhoods (groups of words that only differ in one phoneme) would be more sensitive to this restructuring. In fact, de Cara & Goswami (2003) reported that vocabulary positively influences the effect of phonologic neighbourhood density on the performance of oddity tasks even on children in a pre-literate age. Moreover, not only learning how to read but also schooling (Kosmidis et al., 2006) contribute to vocabulary acquisition. Kosmidis et al., (2006) found a progressively better performance in a lexical decision task – where one must match the heard word with those stored in the auditory input lexicon, which is directly related with the vocabulary size – across illiterates, literates and highly schooled subjects. However, the idea that lexical restructuring depends on literacy or vocabulary is not completely unchallenged. Hogan et al. (2010) report that, in phoneme deletion tasks, the target words neighbourhood density has the same effect in second and fourth graders. Instead, it was word frequency that affected positively fourth graders performance more than second graders. Moreover, Ventura et al. (2006) in two different experiments – the first using a gating task and the second using a word-identification-in-noise-task – reported that word frequency and neighbourhood density affects equally the task performance of illiterates, ex-literates, and literates.

The impact of literacy on the language-processing brain

Literacy acquisition changes the brain in such a permanent and irreversible way that some authors called it a “brainwash” (Blakemore & Frith, 2005; p.72) or a “virus” that affects every area of speech processing (Frith, 1998). Literacy triggers a set of structural and functional changes in the brain which include (but are not limited to) the auditory processing areas, including its connections with 1) cortical and subcortical motor mechanisms that modulate phonology; 2) the multisensory integrative cortex responsible for lexicon-semantic components and 3) the visual cortex and parietal cortical areas responsible for the dorsal visual pathway (Castro-Caldas & Reis, 2000). A study using a word and pseudoword repetition task with literates and illiterates found that, in word repetition, literates activate more the Brodmann area 40 (left supramarginal gyrus) when masking with the activation pattern of words-pseudowords (Castro-Caldas et al., 1998). This area is usually involved in phonological processing (Pattamadilok et al., 2010) and it was hypothesized that it may have a role in an oscillation-based functional model of speech processing (Giraud & Poeppel, 2012). On the other hand, when repeating pseudowords, literates activate different cortical and subcortical areas comparing to repeating words, while illiterates do not show this difference (Castro-Caldas et al., 1998). A subsequent study analysed this dataset using structural equation modelling and found different network interactions between groups in pseudoword repetition including 1) the interaction between the Broca's area and the inferior parietal cortex, 2) the posterior midinsula bridge between Broca's and Wernicke's area and 3) areas comprising the working memory phonological loop (see Baddeley, 1992; 2000 for an overview) network including the link between the anterior cingulate and the auditory cortex (Petersson et al., 2000). Moreover, the activation of inferior parietal areas (including the supramarginal gyrus) in word and pseudoword repetition was more right-lateralized in illiterates while it was more left-lateralized in literates (Petersson et al., 2007).

In a different study, Carreiras et al. (2009) reported lower grey matter density in the medial temporal gyrus (bilaterally), dorsal occipital gyrus (bilaterally) and supramarginal gyrus (left) in illiterates compared to late-literate culturally and age matched. A functional connectivity analysis in the same study showed that reading acquisition increases the functional direct linkage between visual (dorsal occipital gyrus) and phonological (supramarginal gyrus) areas or indirect linkage by means of semantic-related areas (medial temporal gyrus). These authors also found an increase of grey matter density in the left superior temporal gyrus, a structure involved not only in the analysis of the receptive speech spectrotemporal properties but also in the phonemic encoding on speech production (Hickok & Poeppel, 2000; 2007). In fact, temporal areas involved in speech processing are significantly affected by learning how to read. One of those areas is the superior temporal sulcus that is responsible for phonological processing of linguistic units in different timescales (Hickok & Poeppel, 2007). The activation of this structure not only to written sentences but also to spoken language is modulated by reading performance (Dehaene et al, 2010). The activity of the planum temporale (superior temporal region posterior to the Heschl's gyrus) is also modulated by literacy. The planum temporale is activated when listening to speech sounds, showing an increase of activation when there is a successful grapheme-phoneme mapping and a decrease of activation when a grapheme is presented with an incongruent phoneme in normally-developing literate subjects (Atteveldt et al., 2004). However, in dyslexic subjects, the planum temporale is less activated for speech sounds and there is no change in the activation of this area whether a phoneme is presented with a congruent or incongruent grapheme (Blau et al., 2009). In turn, literates activate this area more than illiterates (independently of reading acquisition age) whether when hearing simple sentences or when doing a lexical decision task (Dehaene et al., 2010).

Specific changes in the structure of subcortical features linked to language processing have also been reported. One of those changes comprises the morphology and density of white matter in the corpus callosum. The corpus callosum, specifically the posterior third of this structure – the splenium – is essential for the inter-hemispheric coordination of lexical-semantic and prosodic processing (Sammler et al., 2010). In fact, it was suggested that the corpus callosum may have a role in the transmission of the cortical oscillatory activity information between hemispheres (Boemio et al., 2005). This structure, where parietal fibres are thought to cross, is thinner in illiterates (Castro-Caldas et al., 1999) and the amount of white matter in the splenium of the corpus callosum is positively associated with the ability to read (Carreiras et al., 2009). Moreover, callosal lesions in the splenium as well as in the forceps major and periventricular regions cause pure alexia – a specific loss of reading ability (Damasio & Damasio, 1983). More recently it was also reported that the structure of the arcuate fasciculus improves with literacy (Thiebaut de Schotten et al., 2014). This structure consists in several bundles of axons that link Broca's area, the inferior parietal lobe and the posterior superior temporal lobe and also includes connections with the ventral temporal lobe (Catani & Mesulam, 2008; Thiebaut de Schotten et al., 2014). Lesions on the arcuate fasciculus may cause conduction aphasia – a language disorder affecting naming and repetition – because of the disconnection between anterior (production in Broca's area) and posterior (comprehension in Wernicke's area) language modules (Geschwind, 1965). Lesion studies have stressed the involvement of the arcuate fasciculus in this pattern of language deficits (Damasio & Damasio, 1980) although naming deficits typical of this neuropathology (phonemic errors and phonemic paraphasias) were also reported in subjects with superior-temporal lesions and an intact arcuate fasciculus (see Hickok & Poeppel, 2000). Thiebaut de Schotten et al. (2014) reported that the structure of the arcuate fasciculus is modulated by literacy acquisition (independently of schooling) as the fractional anisotropy of the posterior arcuate is augmented, suggesting a change either in fibre density, myelination, axonal diameter or axon density. The same study also showed a correlation between the degree of activation in the superior temporal area (when spoken sentences were presented) and a region located on the inferior occiptotemporal lobe (when written strings were presented) with the fractional anisotropy of the posterior arcuate.

In this ventral occipitotemporal region, lateral to the mid-portion of the left fusiform gyrus, is found the visual-word-form area (Cohen et al., 2000; McCandliss et al., 2003). This area is initially programmed for the recognition of objects and faces and subsequently (when one learns how to read and write) “recycled” to become attuned to visual word recognition (Cohen & Dehaene, 2004). Moreover, the elicited response by written strings in this area is much stronger in literate individuals than in illiterates (Dehaene et al., 2010) and a case study showed that damage to this area can cause pure alexia (Gaillard et al., 2006). Dysfunction not only in the visual-word-form area but also along the visual pathway attuned to word recognition has been reported on dyslexic children (van der Mark et al., 2009). However, it was suggested that posterior inferior portions of the temporal lobes also have a role in the lexical interface linking phonological and semantic information (Hickok & Poeppel, 2007). In fact, activations near the fusiform gyrus were reported in speech listening tasks even when the recorded speech waveform was reversed, becoming the speech, therefore, unintelligible (Crinion et al., 2003). This data may suggest that, in literates, the visual word representations may be activated during speech processing tasks without the explicit involvement of orthographic material. Indeed, Seidenberg & Tanenhaus (1979), using an auditory rhyme identification task, showed that literates were faster (56 ms) identifying rhymes when they were orthographically similar than when they were orthographically dissimilar. Moreover, literates show higher activation of the visual-word-form area than illiterates in lexical decision tasks (Dehaene et al., 2010). A complete review on how literacy changes the brain and behaviour, including those unrelated with speech processing that fall out of the scope of this work can be found elsewhere (Castro-Caldas & Reis, 2000; Dehaene et al., 2015).

The present study

Recent work has been supportive that the auditory organization of the mental lexicon is strongly related to reading skill, and that difficulties in auditory organization may be the cause of reading backwardness in dyslexia (Leong & Goswami, 2016). As speech to literate adults seems to be dominated by energy in the theta band and tighter phase synchronisation between faster AMs (theta and beta/low gamma), speech to non-literate infants is dominated by energy in the delta band and tighter phase synchronisation between slower AMs (delta and theta), suggesting a difference in the auditory lexicon organization between both groups (Luo & Poeppel, 2007; Leong et al., 2014). Crucially, pre-verbal infants show not only a behavioural preference to listen to motherese (for an overview, see Metsala & Walley, 1998) but also mature neural responses to infant-directed speech but not to adult-directed speech (Peter et al., 2016). While this difference could be due to maturation effects, recent studies show that theta entrainment improves with reading skill (Power et al., 2012). Indeed, literacy changes profoundly the functioning, structure and connectivity of areas not only related to speech comprehension but also involved in phonological processing and encoding in speech production. Previous studies addressing the auditory entrainment and oscillatory activity in privileged timescales have used mostly normally-developing literate children/adults and dyslexic children (Molinaro et al., 2016 also tested dyslexic adults). Crucially, no work about this topic so far has studied normally-developing illiterate subjects (people that never learned how to read and write for purely social reasons). Thus, it is not totally clear if this preferential timescale processing shift is due to maturational processes or to the phonological remapping caused by literacy acquisition.

Given this difference in perceptual organisation between groups that differ in literacy skills, we might expect a corresponding difference in production. One way to address this question is by analysing the temporal modulation structure of literates and illiterates speech. In a recent work, it was developed a new method to analyse speech signal properties – the Spectral-Amplitude Modulation Phase Hierarchy (S-AMPH) model (Leong, 2012; Leong & Goswami, 2015) – focused on the privileged timescales of oscillatory entrainment. This model applies probabilistic amplitude demodulation to retrieve the rhythmic properties of the signal and generates a hierarchical representation of the dominant spectral-temporal modulation patterns in the envelope. The S-AMPH is a low-dimensional spectro-temporal model and uses a specific number of bands and respective bandwidths previously determined using principal component analysis dimensionality reduction of high-dimensional spectral and temporal speech envelopes. The temporal patterning of speech is retrieved across these frequency bands in three different amplitude timescales – delta, theta and beta/low gamma – that form the three-tier AM nested hierarchy (see Leong, 2012 and Leong & Goswami, 2015 for a complete overview). The S-AMPH was used in a recent work to assess the role of delta modulations (detected less efficiently by dyslexics) of spoken words on auditory rhyming awareness tasks (Leong & Goswami, 2016). In the present work, it will be used not only to analyse the temporal structure of literates and illiterates speech but also to compare the rhythmic regularities of slower and faster AMs across groups (similarly to the study of Leong et al., 2014 with infant- and adult-directed speech).

Overall, this study aims to assess the differences in the rhythmic properties of speech production in illiterates and literates. If the shifting from a delta to a predominantly theta rate processing of speech is mediated by literacy, it is expected that the illiterates show higher energy in the delta band and lower energy in the theta band when comparing to literates. Moreover, it is also expected that literates show a tighter phase synchronisation between faster AMs (smaller phonological units) while illiterates show a tighter phase synchronisation between slower AMs (larger phonological units). This study also aims to disentangle the effect of phonological remap in the speech signal that occurs when learning how to read and write from a deeper auditory restructuring due to higher schooling that includes the increase of vocabulary size, reading and phonological skills. This is done by testing 3 distinct groups: a group of illiterate subjects, a group of “low literate” subjects who went to school only to know how to read and write and then started to work and a group of “high literate” subjects who pursued a higher education, having at least 12 years of schooling. As the illiterate population in Portugal is mostly constituted by older adults, the subjects in the literate groups are also, at least, over 69 years-old so that there is no significant age difference between groups. Finally, in order to assess the specific role of the phonology remap in the rhythmical properties of speech, measures of phonological awareness in linguistic units of different grain-sizes (syllable, rhyme and phoneme) and vocabulary are retrieved across the 3 participant groups. This relationship between phonology scores and the signal properties of speech in linguistic nested timescales may shed some insight on the locus of the previously mentioned temporal processing shift.

Methods

Participants

Forty-six older adults aged between 69 and 91 were tested after informed consent (presented in written form for literates and orally for illiterates): 15 subjects were illiterates (2 males; $M_{age} = 80.4$; $SD_{age} = 4.4$), 19 were "low literates" with 4 years of literacy or less (2 males; $M_{age} = 77.9$; $SD_{age} = 5.7$ | $M_{literacy} = 3.3$; $SD_{literacy} = 0.75$) and 12 were "high literates" with more than 12 years of literacy (1 male; $M_{age} = 79.7$; $SD_{age} = 6.3$ | $M_{literacy} = 15.5$; $SD_{literacy} = 2.47$). As age showed an approximately normal distribution across all groups according to Shapiro-Wilk tests (all p's ≥ 0.20) and there was no variance heterogeneity according to the Levéne test ($p = 0.15$), a one-way ANOVA was conducted to verify if there were any age differences across the 3 literacy groups. The result of this analysis revealed the absence of a main effect of age across groups [$F(2, 43) = 0.97$; $p = 0.39$] and subsequent mean comparisons between each group showed no significant differences in age (all p's > 0.18). In summary, the participant groups in this study differed in schooling but not in age.

Some precautions were taken to consider a subject completely illiterate. Previously to the experiment, it was assessed if the subject had attended school as a child or had any literacy training as an adult. Any of these would lead to the exclusion of the subject. Further, a list of monosyllabic and disyllabic words was presented (e.g. *cão, gato, casa, "*dog", "cat", "house" in English). The successful reading of any of this target words would exclude the subject from the illiterate group. All participants reported normal or corrected to normal eyesight and no history of neurological disease previous to the start of the experiment. On the other hand, for the literate participants, it was assessed if they had started to attend school at childhood or adulthood. Only the participants that learned to read and write in childhood were included in this study, since the impact of literacy in the brain differs with age of acquisition (see Castro-

Caldas et al., 2009). Because the participants were older adults, a self-report questionnaire of activities of daily living and a cognitive screening test were administered to rule out possible cognitive decline or dementia. Therefore, a scale of instrumental activities of daily living (IADL-B; Lawton & Brody, 1969 adapted to Portuguese by Madureira et al., 2008) and the mini-mental state examination (MMSE; Folstein, Folstein & McHugh, 1975 adapted to Portuguese by Guerreiro et al., 1994) were administered. The self-reported incapacity to perform more than one of the daily activities present in the IADL-B as well as a score in the MMSE below the cut point (Guerreiro et al., 1994) would exclude the participant from this study. The mean score in the MMSE was 22.9 for illiterates, 26.9 for low literates and 29.6 for high literates.

Experimental set-up, tasks and materials

All the testing sessions took place in previously designated quiet rooms (usually small offices or living rooms) with soft furnishings and the subject would sit at a table in the middle of the room, away from walls. The design of this study comprised two different experimental phases. The first phase consisted in the acquisition of phonological awareness and vocabulary measures. For this purpose, the participants performed the following tasks:

<u>Syllabic division</u>: Adaptation of the paradigm of Lieberman et al. (1974) used for the assessment of syllabic awareness in pre-literate children. Under the guise of a tapping game, the participant was asked to segment a word by tapping in the table while repeating it and then saying how many “parts” the word had. The subject would receive the instruction until complete understanding was verbalized. Moreover, feedback was provided for the first two items of the task. The 24 target words used on this task (8 monosyllable, 8 disyllable and 8 trisyllable words) were very frequent ($M_{LogFrequency} = 3.54$; $SD_{LogFrequency} = 0.38$) European Portuguese words and contained V, CV, VC, CVC, CCV and CCVC syllables. Each correct item was scored with 1 point and each error with 0 points in a total score of 24 points.

<u>Rhyme detection</u>: This was similar to the PALPA-P 15b subtest (Kay, Ruth & Coltheart, 1996; adapted to European Portuguese by Castro, Caló & Gomes, 2007) with different target word-pairs. In this forced-choice task, the experimenter would read a

word-pair and would ask the subject to say if the pair was or was not constituted by rhyming words. For subjects who were not familiarized with rhyming, a simple explanation (words that had the same ending, like in a poem or a song) and examples of rhyming and non-rhyming pairs (different from the ones in the task) were provided. Feedback in the first 2 items was also provided. For this task, 3 lists of 12 words – target, rhyme and non-rhyme – were selected. Therefore each target was presented twice, once with a rhyming and once with a non-rhyming pair (e.g. *cedo-dedo* and *dedo-dado*). The words were selected so that the targets, rhymes and non-rhymes would not differ in number of phonological and orthographic neighbours, number of letters and frequency (all F's < 1). The rhymes and non-rhymes only differed in one (minimal pairs) or two phonemes from the targets so that the rhyme judgment could not be performed based on general phonological similarity. Each correct item was scored with 1 point and each error with 0 points in a total score of 24 points.

Phoneme deletion: Similar to the one conducted by Morais et al. (1979), except that the target stimuli were only words and no specific list of practice trials was used before the experimental phase. In this task, the subject is asked to delete the initial sound of a word provided orally by the experimenter. As previously, instructions and examples were provided before starting the experimental task and feedback for the 2 first items was provided. As illiterates have consistently showed to be unable to perform this type of tasks (e.g. Morais et al., 1979, 1986), the task was interrupted after 5 consecutive errors to avoid an excessive decrease in motivation. The target words were created so that phonological neighbourhood and frequency would be manipulated orthogonally (as in Ventura, 2006): half of the high-frequency ($M_{LogFrequency} = 3.73$; $SD_{LogFrequency} = 0.53$) and low-frequency ($M_{LogFrequency} = 1.61$; $SD_{LogFrequency} = 0.50$) words were from sparse neighbourhoods ($M_{Neighbourhood} = 1.94$; $SD_{Neighbourhood} = 1.44$), and half from dense neighbourhoods ($M_{Neighbourhood} = 11.31$; $SD_{Neighbourhood} = 3.7$). Therefore, 4 lists of 8 words that did not differ in syllable number (monosyllables and disyllables) or in the total number of letters or phonemes (all ANOVAs and post-hoc tests p's > 0.23) were created. As expected, one-way ANOVAs revealed a main effect of frequency [$F(3,28) = 45$; $p < 0.001$] and phonological neighbourhood density [$F(3,28) = 33.7$; $p < 0.001$] across lists. Post-hoc comparisons showed that the 2 high frequency lists did not differ (independently of the neighbourhood density, all p's > 0.65) and the 2 dense neighbourhood lists did not differ in density (independently of frequency, all p's >

0.23). The same pattern was observed in low frequency and sparse neighbourhood lists. Density differed significantly between dense neighbourhood lists and sparse neighbourhood lists (all p's < 0.001) and frequency differed significantly between high frequency lists and low frequency lists (all p's < 0.001). In some words, the deletion of the first phoneme would result on a different word while in others would result in a pseudoword. Just like in Morais et al. (1986), in certain words, the first phoneme (a vowel) would also correspond to the first syllable of the word. Each correct item was scored with 1 point and each error with 0 points in a total score of 32 points.

For every phonological awareness task, the estimation of word frequency was based on a European Portuguese corpus containing written (15,354,243) and spoken (856,195) tokens (LMCPC: http://www.clul.ul.pt). The distribution of word frequencies was normalised using a log transformation of raw frequency data. On the other hand, orthographic and phonological neighbourhood densities were estimated based on the *Porlex V2* database (Gomes, Castro & Lima, 2005) that contains 27,374 word entries.

Vocabulary: For a measure of vocabulary, it was administered the vocabulary subtest from the WAIS-III (Wechsler, 1997; translated and adapted to European Portuguese by Barreto, Moreira & Ferreira, 2008).

After this first assessment it would start the second experimental phase, where speech samples were acquired from each subject. To acquire samples of spontaneous speech, it was conducted and recorded a semi-structured interview with the duration of 10 – 15 minutes. To record the subject's speech, a cardioid microphone (Audio-Technica ATR2100-USB) was used, held steady in a microphone stand in a close but comfortable distance to the subject. The microphone was connected via USB to a portable computer with the software Audacity (http://audacityteam.org/about/). The speech samples were digitally recorded at 44.1 kHz. The interview topics chosen were the ones that, as previously identified by Davidson et al. (2003), engaged a larger number of older adults in a conversation. Topics like health issues or other potential emotionally-disturbing topics were avoided to prevent unwanted sources of variability of the speech rhythmic properties. Therefore, 3 main topics were chosen: 1) past professional career life experiences; 2) regular weekly activities (hobbies and routines) and 3) family, friends and neighbours (who still keep in contact with the subject). The elicitation cues were introduced by the experimenter using the following open-ended question: “Tell me about your [topic]” (as in Orange, Kertesz & Peacock, 1998 and Brady, McKenzie & Armstrong, 2003). The criteria to move to the following topic was the same as in Brady, Armstrong & Mackenzie (2006): the participant made a topic closing statement such as “that’s it” or “that’s everything” followed by silence (> 3 s), stopped speaking (> 3 s) or indicated in another manner that the topic or his/her turn was complete.

Although the model of speech rhythm analysis used in this thesis (see next section) was validated in English (see Leong, 2012), no study so far has tested its applicability in the analysis of European Portuguese utterances. Crucially, it would be necessary to test the sensitivity of this model to speech rhythm. Therefore, an additional task – proverb repetition – was created to acquire speech samples with enhanced rhythmic properties. Portuguese proverbs, much like English proverbs, have specific rhythmic properties granted by the use of alliteration, assonance, consonance and rhymes (see Zhang, 2012 for an overview). To guarantee even further the rhythmicality of the proverbs, these were carefully selected (from the proverb book of Estanqueiro, 1996) so they could match a specific metrical syllabic pattern (like the nursery rhymes in Leong & Goswami, 2015) and so they could be trochees (strong-weak syllable alternation as in “*Muito riso pouco siso*”, strong syllables underlined), iambs (weak-strong syllable alternation “*Galinha velha faz bom caldo*”) or dactyls (strong-weak-weak syllable alternation “*Vozes de burro não chegam ao céu*”). A total of 6 proverbs (2 of each rhythmic structure) were repeated by the subjects. If the analysis used is, in fact, sensitive to the specific rhythm of linguistic units, it is expected an enhancement in the rhythmic regularity measure used (see below).

Data processing & analysis

Before being analysed, the recordings were manually divided into shorter segments using Audacity software. As in Leong et al. (2014), each spontaneous segment contained a complete phrase from an original speech utterance. On the other hand, rhythmic speech segments contained a full proverb. Only the most noise-free segments containing speech sounds from the participant were chosen for the analysis. On what concerns to spontaneous speech samples, the illiterates contributed with 89 segments, the low literates with 107 segments and the high literates with 90 segments in a total of

286 speech segments. These segments ranged from ~4s to ~20s and each subject contributed, on average, with 6.2 segments (range 3 – 11). Each participant also contributed with 6 rhythmic speech utterances ranging from 1s to 3s each. Each speech segment was z-scored to standardise its mean and standard deviation and the S-AMPH model of its amplitude was extracted using a 2-stage filtering process (Leong, 2012; Leong & Goswami, 2015). In this process, the z-scored acoustic signal was band-pass filtered into five frequency bands (channel edge frequencies: 100, 300, 700, 1750, 3900 and 7250 Hz) using a series of adjacent finite impulse response (FIR) filters. Next, the Hilbert envelope was extracted from each band-filtered signal and these five Hilbert envelopes were down-sampled to 1050 Hz and passed through a second series of band-pass filters in order to isolate the three different AM bands within the envelope modulation spectrum. These three AM bands corresponded to delta rate modulations (0.9-2.5 Hz), theta rate modulations (2.5-12 Hz) and beta/low gamma rate modulations respectively (12-40 Hz). The result of this two-step filtering process was a 5 (frequency) x 3 (rate) spectro-temporal representation of the speech envelope, described as comprising delta-, theta- and beta/low gamma-rate AM bands for each of the 5 spectral bands (Leong & Goswami, 2015).

To assess the temporal structure of the speech signal, two measures were derived from the S-AMPH model in this work: the envelope modulation spectrum and a multi-timescale synchronization index (as in Leong et al., 2014). To calculate the modulation spectrum of the speech samples, the sub-band Hilbert envelopes of the stimuli (resulting from the previously-described S-AMPH decomposition procedure) were individually passed through a modulation FIR filterbank with 24 channels logarithmically-spaced between 0.9-40 Hz. For each speech sample, and each frequency sub-band, the mean power across all modulation channels was computed, and the relative power difference from this mean was computed for each modulation channel. The differenced modulation power spectrum was then averaged across the 5 frequency sub-bands for each speech sample of every subject across the three literacy groups. The 24 points in the modulation spectrum were divided up to reflect the Delta , Theta, Beta/low gamma band divisions used in the SAMPH and as an approximate reflection of the neural oscillation bands described in the multi time resolution model of speech processing (Poeppel 2003; Poeppel, 2014). To assess the differences between the signal's energy across these bands and literacy groups, the area under the curve corresponding to the three modulation bands was computed and a 3x3 (bands by literacy group) repeated measures ANOVA was conducted.

Subsequently, a measure of multi-timescale phase-synchronisation, the Phase Synchronisation Index (PSI) was computed between nested modulation rate bands in the S-AMPH representation of each speech segment (i.e., delta-theta phase synchronisation and theta-beta/low gamma phase synchronisation). PSI computation was performed separately for each of the 5 spectral bands in the S-AMPH representation. The *n:m* PSI was originally conceptualised to quantify phase-synchronisation between two oscillators of different frequencies (e.g. muscle activity; Tass et al, 1998), and was then adapted for use in neural analyses of oscillatory phase-locking (e.g. Schack & Weiss, 2005). This adaptation is applied to speech AMs in the S-AMPH model. The PSI was computed as:

$$\mathrm{PSI} = \left| \left\langle e^{i(n\theta_1 - m\theta_2)} \right\rangle \right|$$

In this equation, *n* and *m* are integers describing the frequency relationship between the two AMs being compared. As in Leong et al. (2014) for the delta-theta AM band analysis, an *n:m* ratio of 2:1 was used, while for the theta-beta/low gamma AM band analysis, an *n:m* ratio of 3:1 was used. The values θ1 and θ2 refer to the instantaneous phase of the two AMs at each point in time. Therefore, (nθ1 - mθ2) is the generalised phase difference between the two AMs, which was computed by taking the circular distance (modulus 2π) between the two instantaneous phase angles. The angled brackets denote averaging of this phase difference over all time-points. The PSI is the absolute value of this average, and can vary between 0 (no synchronization, typical of a sound random in rhythm) and 1 (perfect synchronization, typical of a sound with perfect rhythmic regularity). Delta-theta and theta-beta/low-gamma PSIs across the 5 spectral bands were averaged in every speech segment. Further, individual PSI scores were estimated for every subject across all three groups. For that purpose, a grand mean of the PSI values across speech segments was calculated for each participant. Therefore, it was assigned to every subject two synchronisation scores: a delta-theta PSI and a theta-beta/low-gamma PSI score. Shapiro-Wilk tests confirmed the normal distribution of both PSI scores (all p's > 0.14) and Levéne tests confirmed variance homogeneity (all p's > 0.11) across the 3 groups in spontaneous and rhythmic speech. Box M tests showed no differences in the covariate matrix between rhythmic and spontaneous speech in both PSIs in slower and faster AM pairs. Therefore two one-way MANOVAs were conducted for each AM-pair (delta-theta and theta-beta/low gamma) to analyse the data.

The relationship between literacy, phonology and vocabulary with the rhythmic properties (PSIs) of the spontaneous speech were analysed by calculating the Pearson correlations between these variables. Although there was no age difference between the 3 groups, since there was considerable variance in age among the participants in each group (the younger participant was aged 69 and the older participant was aged 91) and age has been shown to influence neuropsychological test performance, all correlations controlled for age (partial correlations). The extraction of the segment's amplitude by the S-AMPH model as well as the calculation of the envelope modulation spectrum and PSI values were performed using MATLAB R2013a. The PSI comparisons as well as the correlation coefficients were calculated using the STATISTICA software.

Results

Phonological awareness and vocabulary

The mean score performance in the behavioural measures assessed in the present study can be found in Table 1. The vocabulary score presented in the table is the raw score calculated according to the WAIS-III manual. The overall score for phonological awareness was calculated as the sum of rhyme detection, phoneme deletion and syllabic division scores. Since phonological awareness measures did not reveal an approximately Gaussian distribution ($p < 0.05$ using Shapiro-Wilk tests), non-parametric tests were conducted to compare the differences in performance between groups. Kruskal-Wallis tests showed, in every assessed measure, a very significant effect of group (all p's < 0.0001). Subsequent Mann-Whitney tests showed that high literates performed better than low literates and that low literates performed better than illiterates in vocabulary and phonology tasks (all p's < 0.013). Therefore, as expected, vocabulary and phonology scores increase with schooling and literacy in this sample. Notably, although phonemic awareness is developed while learning how to read (Ziegler & Goswami, 2005) and a poor performance would be expected on phoneme deletion tasks in illiterates (Morais et al., 1976, 1986), a floor effect is reported in this work. This could be due to (1) the feedback being only given in the first 2 items and illiterates may need more practice trials to get used to this task (in Morais et al., 1979 there was a pre-test phase consisting on 15 introductory trials); (2) the first phoneme of the target words varying widely in the present study (just like in Loureiro et al, 2004 who also reported very poor performances in illiterates on this task) and that could decrease the generalizability of the feedback given in the first trials (in Morais et al., 1979 it was always the same phoneme that should be deleted); (3) the illiterates in this study had never been to school while in Morais et al. (1979) illiterates could have up to 6 months of schooling possibly leading to a slightly more developed phonological awareness; (4) normal age-related executive functioning decline (Fisk & Sharp, 2004) may render this metalinguistic task (considered an executive function task in Ardila et al., 2010) impossible to perform in older adult subjects like the ones in this study.

Table 1: Mean performance (standard deviation in brackets) on the phonological awareness and vocabulary tests. Every measure improves across literacy groups: illiterates < low literates < high literates

	Illiterates (n = 15)	**Low Literates** (n = 19)	**High Literates** (n = 12)
Phonological awareness (composite)	23.8 (4.7)	58.4 (11.6)	74.8 (5.1)
Syllabic division	11 (3.4)	19.7 (4.3)	23.8 (0.5)
Rhyme detection	12.8 (3.2)	16.1 (3.8)	21 (3.1)
Phoneme deletion	0 (0)	22.6 (6)	30.1 (2.4)
Vocabulary (WAIS-III)	9.7 (3.6)	26.5 (11.5)	54.8 (5.7)

Modulation spectra

Figures 1-4 plot the speech wholeband envelopes and modulation spectra for each one of the three literacy groups averaged over the 5 acoustic frequency bands of the S-AMPH model. Fig. 2 shows that rhythmic speech is characterised by two peak (highest power) modulation rates. The first peak corresponds to modulation rates in the delta band (~1.89 Hz for high literates and ~2.21 Hz for low literates and illiterates) and the second peak corresponds to modulation rates in the theta band (~3.53 Hz for low literates and ~4.13 Hz for illiterates and high literates). Indeed, this is in line with the fact that rhythmic utterances (particularly in the iambic and trochaic patterns) had a very similar number of stressed/unstressed syllables (Fig.1). This pattern was not modulated by literacy since it was consistent across all groups. On the other hand, as expected, spontaneous speech is characterised by a less rhythmical alternation between stressed and unstressed syllables (Fig.3). Similarly to previous findings (Greenberg et al., 2003; Giraud & Poeppel, 2012; Ding et al., 2016; Leong et al., 2014), a single peak in the theta range (~4.13 Hz for the three groups, see Fig.4) was found for spontaneous speech. The mean modulation energy values for the different bands in spontaneous speech across groups are represented in Fig.5. On average, illiterates seemed to show higher energy in the delta band and lower energy in the theta band. However, the repeated measures ANOVA with the energy in each band as the dependent variable revealed no significant effect of group [F(2,43) = .734, p = .486] or group x band interaction [F(2,43) = 1.547, p = .224]. Overall, there is a predominance of stress- and syllable-rate modulations in rhythmic speech, while spontaneous speech is predominantly characterised by the syllable rate. Moreover, no differences were found between literacy groups in the modulation energy of delta theta or beta/low gamma AMs.

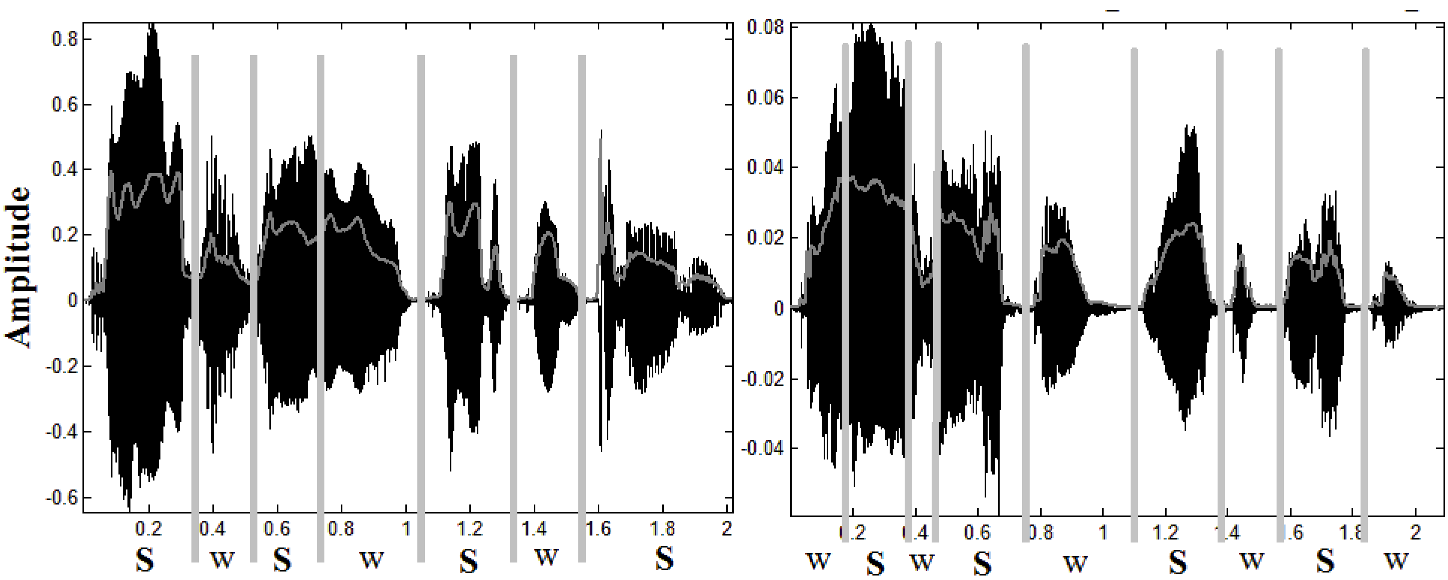


Fig.1: Example plots of the structure (signal amplitude across time in seconds and wholeband envelopes) of rhythmic utterances with trochaic (left) and iambic (right) patterns. The vertical grey stripes denote syllabic boundaries. S = stressed syllable; w = unstressed (weak) syllable.

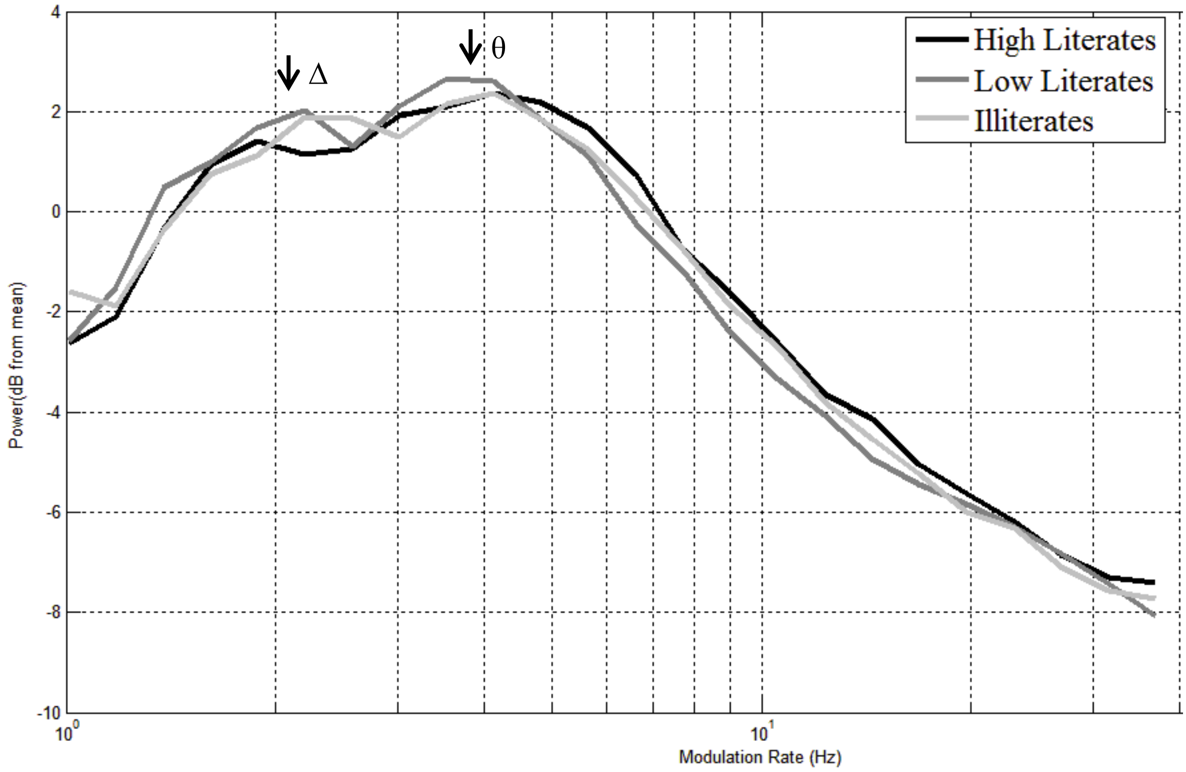


Fig.2: Grand mean modulation spectra averaged across frequency bands and speakers from the different literacy groups for rhythmic speech. The x-axis shows the modulation rate, the y-axis shows the power in dB (normalised according to the mean power of each frequency band in each sample). Arrows represent the delta and theta peaks.

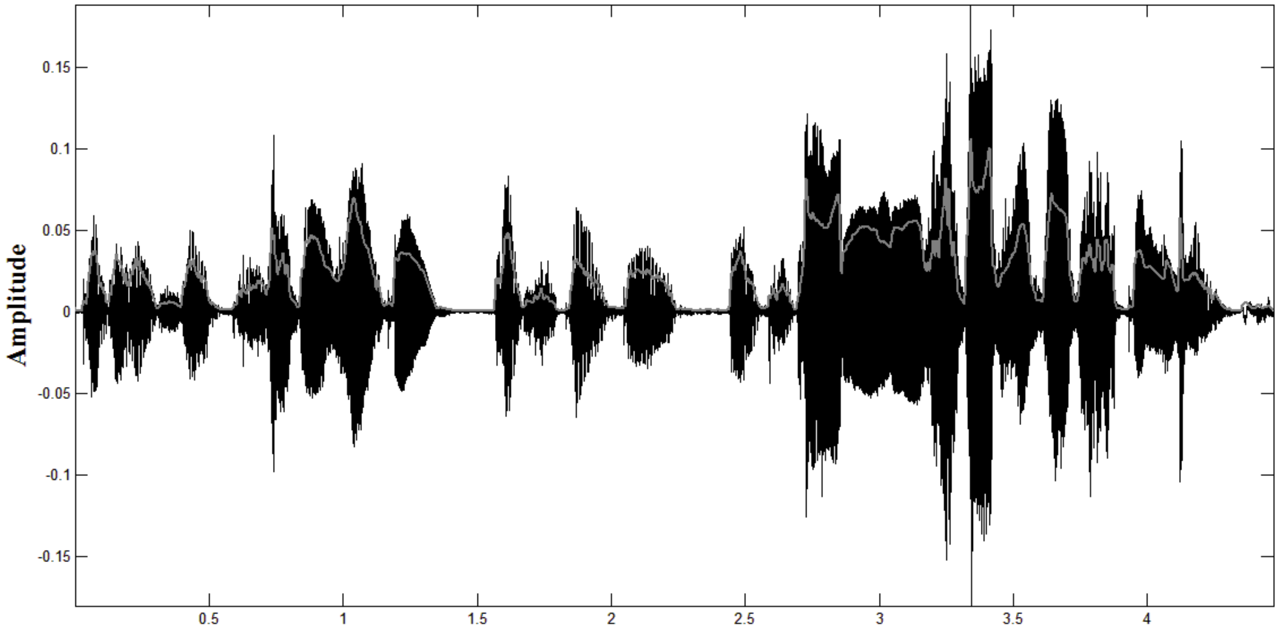


Fig.3: Example plot of the structure (signal amplitude across time in seconds and wholeband envelope) of a spontaneous utterance from a literate subject.

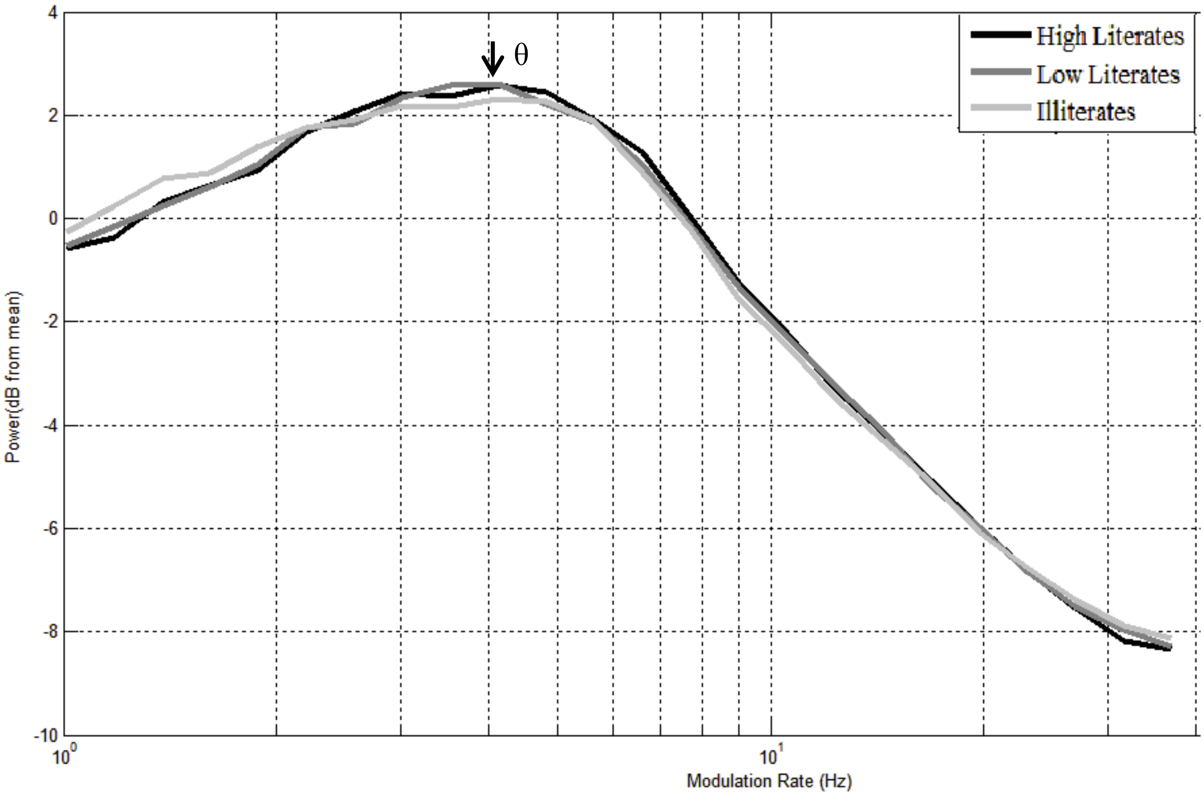
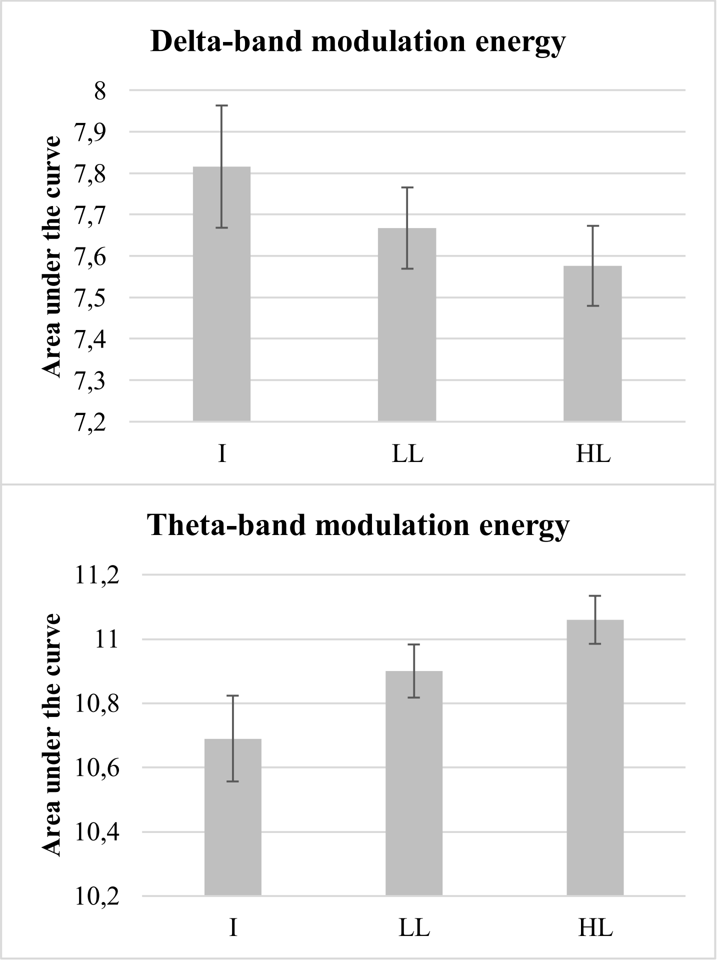


Fig.4: Grand mean modulation spectra for spontaneous speech. The arrow marks the theta AM peak.



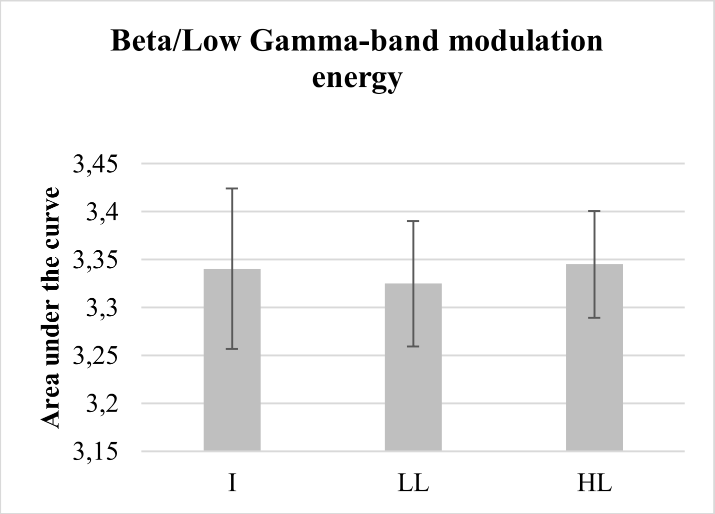


Fig.5: Modulation energy for spontaneous speech in the 3 AM bands across the 3 literacy groups. I = Illiterates; LL = Low Literates; HL = High Literates.

Phase synchronisation

Fig.6 shows the mean PSIs for rhythmic and spontaneous speech across literacy groups in the two AM pairs. The PSI measures showed to be sensitive to the rhythm of speech units, since the rhythmic speech had significantly higher synchronisation values than spontaneous speech in both AM-pairs [delta-theta PSI: $t(45) = 21.77$; $p <<.001$ | theta-beta/low gamma PSI: $t(45) = 15.35$; $p <<.001$]. Therefore, our results confirm that the PSI measure calculated on the low-dimensional representation of the envelope provided by the S-AMPH is sensitive to the rhythmic changes of speech in European Portuguese subjects.

Furthermore, if literacy affects the temporal structure of speech, differences across groups are expected in spontaneous speech. On the other hand, if the rhythmic properties of the utterances are controlled, it is expected that all groups show similar PSI values, even if literacy influences the produced speech rhythm. MANOVAs (F1 for delta-theta and F2 for theta-beta/low gamma PSIs) showed an effect of literacy group for both delta-theta [$F_1(4,84) = 3.71$; $p = .008$; Wilk's $\Lambda = .722$; partial $\eta^2 = .15$] and theta-beta/low gamma PSIs [$F(4,84) = 2.822$; $p = .03$; Wilk's $\Lambda = .78$; partial $\eta^2 = .12$]. Post-hoc Tukey tests revealed no differences between groups in rhythmic speech for both AM pairs (all p's > .26). On the other hand, for spontaneous speech, illiterates showed lower delta-theta PSI values when compared to low literates ($p = .039$) and high

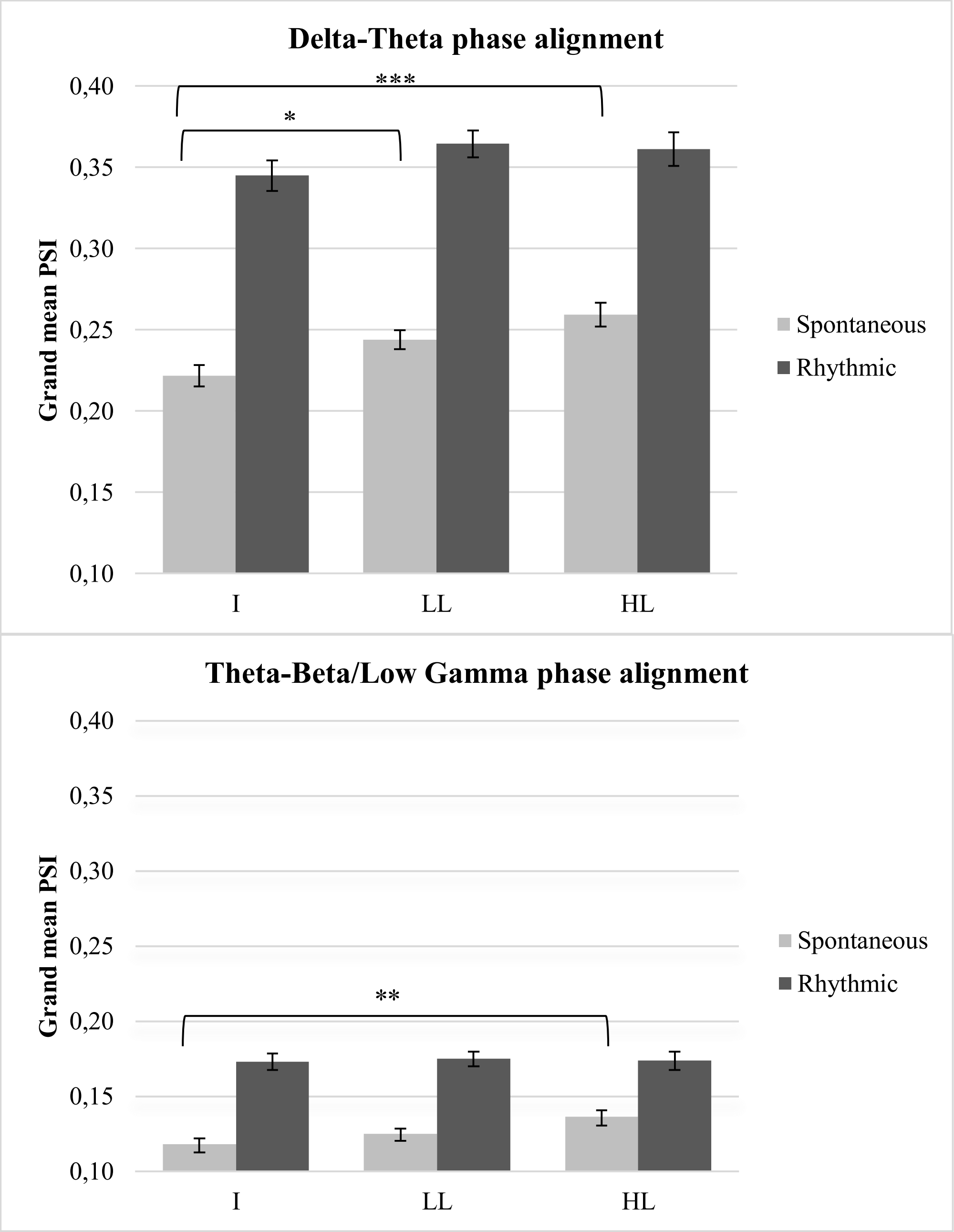


Fig.6: Phase synchronisation in rhythmic and spontaneous speech across the 3 literacy groups. I = Illiterates; LL = Low Literates; HL = High Literates. *p < .05; ** p < .01; ***p < .001.

literates (p = .001). No differences were found when comparing low literates with high literates (p = .24). These results were confirmed using a Kruskal-Wallis ANOVA and post-hoc Mann-Whitney tests. Overall, the synchronisation between AMs corresponding to larger phonological units (stress and syllable) in spontaneous speech was affected by literacy acquisition. Finally, for theta-beta/low-gamma PSI scores, post-hoc Tukey tests revealed significant differences between high literates and illiterates (p = .006) but no significant differences between illiterates and low literates (p = .35) or between low literates and high literates (p = .1). These results were also confirmed using non-parametric tests and a post-hoc Mann-Whitney test suggested that low literates and high literates differed significantly in theta-beta/low gamma PSI scores (U = 58; Z = -2.25; p = 0.024). Overall, the synchronisation between AMs corresponding to smaller phonological units (syllable and onset-rhyme/phoneme) was affected by literacy but in a smaller extent when comparing to larger phonological units.

Behavioural assessment measures and spontaneous speech synchronisation

Since the energy across bands was not significantly different across groups, only the partial correlations between the performance in the phonology / vocabulary assessments and PSI scores are shown in Table 2. Partial correlations between years of literacy and PSI scores were also calculated. In both amplitude modulation PSI scores, positive moderately strong correlations with the behavioural measures and schooling were observed. Phonology measures (namely phoneme deletion and syllable division tasks) showed a slightly higher correlation with PSI scores between larger phonological units than with scores between smaller phonological units.

Table 2: Partial correlations between behavioural measures / literacy and the PSI scores (controlled variable – age).

	Delta-Theta PSI	**Theta-Beta/Low gamma PSI**
Literacy (years)	r = .45; p = .002	r = .44; p = .002
Phonological awareness (composite)	r = .55; p < .001	r = .41; p = .006
Syllabic division	r = .51; p < .001	r = .36; p = .014
Rhyme detection	r = .53; p < .001	r = .49; p = .001
Phoneme deletion	r = .50; p < .001	r = .38; p = .009
Vocabulary (WAIS-III)	r = .42; p = .004	r = .47; p = .001

Discussion

The present work aimed to study the effect of literacy on the speech temporal modulation structure. Specifically, the variance in the amount of energy and phase alignment between delta, theta and beta/low gamma AMs (corresponding to stress, syllable and phoneme rates) of spontaneous speech across literacy groups (illiterates, low literates and high literates) was assessed. The sensitivity of the speech analysis method (S-AMPH) to the rhythmic properties of speech units in European Portuguese was also assessed and confirmed by analysing evoked rhythmic speech from the participants of all groups. It was found that the phase-alignment of nested AMs of speech units while repeating well-known rhythmic sequences of words (proverbs) was significantly higher when comparing to spontaneous speech and was not affected by literacy. This not only confirms the sensitivity of the S-AMPH to the rhythmic properties of European Portuguese speech as it also suggests that when a rhythmic framework for speech is provided beforehand, illiterates and literates alike can produce speech with the same temporal structure. For spontaneous speech, although illiterates showed a tendency for higher energy in the delta band and lower energy in the theta band, such differences were not statistically significant, which does not support the hypothesis that energy in the delta and theta bands in speech production would be modulated by literacy.

On the other hand, the rhythmical regularities of nested AMs in spontaneous speech vary significantly with literacy. However, unlike our previous hypothesis, illiterates showed a consistently lower phase alignment not only across faster but also across slower nested AMs when compared to literates. Specifically, illiterates showed significantly lower phase synchronization between delta- and theta-rate AMs (stress and syllable) when compared to low literates and high literates. Moreover, illiterates showed lower phase synchronization between theta- and beta/low gamma-rate modulations (syllable and onset-rime/phoneme) compared to high literates but not compared to literates, which may represent a general effect of literacy not confined to the explicit awareness of phonemic units. Both PSI values between faster and slower modulations showed a moderately strong positive correlation with vocabulary, phonology measures and years of literacy.

The reported differences in phase alignment between the speech units AMs in illiterates compared to literates suggest that the literacy and phonology development modulate the cognition of word retrieving. As stress and syllable AMs show a literacy-induced enhancement, this could be due to the deeper restructuring of the lexicon and phonology remap that takes place when learning to read. Metsala & Walley (1998) suggest that prosodic structure and overall acoustic shapes of words may actually be the first lexical representations available for infants (9 – 18 months of age) and that the development of syllable sensitivity also happens early in language development. Lindfield, Wingfield & Goodglass (1999) suggested that syllabic stress is also represented in the mental lexicon of adults after reporting a facilitation effect of word recognition in a gating task when the words' prosodic pattern was provided. However, literacy skills may play a part of prosodic stress processing. Goswami et al. (2016) found that although typically developing children and dyslexics performance in a short-term memory task was affected by the target words prosodic similarity, dyslexics recalled fewer words. Moreover, in the same study, stress sensitivity, vocabulary and phonological awareness were the best predictors of the participants' performance in the memory task. Therefore, lexical restructuring may not only refine the syllabic- but also the stress-based representation of words and the present work supports that this refinement affects significantly the spontaneous speech signal. In fact, Webster & Plante (1995) had already reported that rather than age, it is phonological awareness that is more related to the phonology of speech production in young developing children.

Literacy effect was smaller in the phase alignment between syllables and phonemes. This is a curious result since one of the hallmarks of literacy is the acquisition and training of phoneme awareness (Ziegler & Goswami, 2005; Anthony & Francis, 2015) and since areas typically recruited in phoneme processing in speech perception are some of the most influenced by literacy (e.g. supramarginal gyrus and planum temporale, see Literature Review section). However, in the present work, only high literates showed a consistent enhancement in the synchronisation of these AMs when compared to the illiterate group. Therefore, a main question arises: why don't phonemic units show a stronger literacy-induced modulation?

Literacy as a potential modulator of word-form encoding and monitoring

A way to further understand this difference in the rhythmic properties of illiterates and literates speech is to analyse the computational neuroanatomy of speech production. Hickok (2012), integrating two different approaches to the study of speech production (namely motor control and psycholinguistics), proposed the Hierarchical State Feedback Control (HSFC) model. The input to this model starts with the activation of a conceptual system that, in turn, activates a specific lemma (word). Subsequently, from the word level, it takes place the projection to a high-level feedback control, where the auditory syllable targets (activated by the superior temporal regions) and the motor syllable programs (controlled by the Brodmann area 44) interact through the Sylvian parietal-temporal area. Further, this high-level loop projects to a lower-level loop where somato phoneme targets (controlled by the anterior supramarginal gyrus and the primary somatosensory cortex) and motor phoneme programs (controlled by the ventral portion of the Brodmann area 6 and primary motor cortex) interact through the cerebellum (Hickok, 2012).

While the neuroanatomic areas implicated on motor processing in the HSFC do not seem to be structurally affected by literacy, the areas supporting the phonological processes are significantly modulated by literacy. Indeed, the areas supporting the phonologic hierarchical processes (exception made to the primary somatosensory cortex) show an increase on grey matter density in literates when compared to illiterates (Carreiras et al., 2009). This difference is most visible in areas that control syllable targets according to the HSFC, namely the superior temporal sulcus and superior temporal gyrus. In a review of functional imaging literature of speech production, Indefrey and Levelt report that the left posterior superior / middle temporal gyri are activated in the retrieval of the phonological code of each selected lemma (Indefrey & Levelt, 2004). Thus, the difference found in the speech rhythmicality between literates and illiterates may be due to a difference in the development of phonological encoding areas active in speech production. Moreover, the same authors posit that the superior temporal gyrus monitors the output of the syllabification process in speech production. Syllabification consists in assigning the segments of the morphemes to a metrical structure, determining, according to the phonological context of the utterance, the number of syllables and main stress position of the produced words (Levelt, Roelofs & Mayer, 1999). Interestingly, in the present study, the PSIs between AMs corresponding to stress and syllable units were the ones that increased mostly with literacy.

According to Hickok (2012, 2014), as the word is encoded at a higher (auditory and motor) level as a syllable or a metrical sequence or syllables, the sequence of phonemes (lower level processing) is built subsequently in to this frame. Moreover, in this process where the word encoding shifts hierarchically from syllable to phoneme targets (articulatory feature clusters), somatosensory and motor areas which are not affected by literacy have a crucial role (Hickok, 2012). Phoneme encoding in speech production also recruits, however, part of the supramarginal gyrus (Hickok, 2012, 2014) and as it was previously mentioned, this region is functionally and structurally modulated by literacy (Castro-Caldas et al, 1998; Petersson et al., 2007; Carreiras et al., 2009) being also implicated on phonological processing in speech perception (Pattamadilok et al., 2010). While phonemic encoding in speech production only partially recruits literacy-modulated areas, in the present study we have also found a smaller variance in the theta-beta/low gamma (syllable and onset-rhyme / phoneme units) PSI scores across literacy groups and only literates with highly developed phonology had significantly higher scores when comparing to illiterates. Nevertheless, the tendency of the rhythmic regularity pattern in spontaneous speech, even between these smaller units, was illiterates < low literates < high literates.

Overall, literacy and phonology show stronger effects on the phase-locking of units highly dependent on literacy-modulated areas for their encoding. On the other hand, linguistic units less dependent on these literacy-modulated areas in speech production show a smaller effect of literacy and phonology on their PSI scores. While in speech perception, a group of auditory processing areas are involved in phonemic processing (Atteveldt et al., 2004; Blau et al., 2009), in speech production, phonemes are processed mostly in primary motor and somatosensory areas as well as the cerebellum. According to Hickok (2012), this phoneme processing depends on the access of a mental syllabary where the phonetic codes and gestural scores of syllables are stored. The access to the syllabary is modulated by the syllable frequency (even for pseudowords, see Cholin, Levelt & Schiller, 2006) but not by the complexity of the phonemic codes of a syllable (Levelt & Wheeldon, 1994; Levelt, Roelofs & Mayer, 1999). Therefore the access to phonemes in speech production may depend more on spoken language experience and development than on literacy. Previous work had already suggested that phoneme discrimination in speech production develops with age. Nittrouer, Studdert-Kennedy & McGowan (1989) found that adults discriminated phonemes (fricatives) better than older children (7 year-olds) and older children discriminated phonemes better than younger children (3-5 year-olds) in a speech production task. On the other hand, the present study controlled for developmental variables (only older adults were tested) and no literacy effect was found. However, we present some evidence that not only age but also highly developed literacy skills can influence phoneme processing, namely how they are spaced in the syllabic structures, suggesting that a continuous development of phonologic skills has influence on the speech rhythmicality.

Beyond literacy: enhancing phonology through music

While this thesis provides evidence that literacy and phonology skills can modulate the rhythmicality of spontaneous speech, previous work has also suggested that the improvement of timing abilities in rhythm production due to music training may generalize to non-musical (i.e. musically implausible as in Cameron & Grahn, 2014) rhythm. In fact, in an ERP study, musicians showed higher sensitivity to syllabic temporal structure – larger P200 – and a stronger reanalysis of the rule-based expectancy of the metrical structure with a larger P600 response for metrically incongruous words (Marie, Magne & Besson, 2010). Music training is, indeed, an external factor that can change the brain's cortical and subcortical structures. One of those structures is the corpus callosum, although unlike literacy it is not the posterior but the anterior portion that is modulated and shows a larger volume (Schlaug, 2001; Schlaug et al., 1995) as well as a change in its white matter architecture (Schmithorst & Wilke, 2002) when comparing musicians with non-musicians. Other studies like Schneider et al. (2002) found an increase of grey matter on the Heschl's gyrus bilaterally in professional musicians and Schlaug (2001) report an increased left-sided PT asymmetry in musicians with absolute pitch.

On the other hand, functional imaging studies report that musicians show enhanced phase-locking (suggesting the auditory cortex is better able to represent sound) and larger gamma-band responses to pure and musical tones (Trainor, Shahin & Roberts, 2009). Most importantly, music training improves sensory-perceptive skills like sensitivity to frequency, duration and intensity, which are the same parameters on which speech processing relies on (for a review see Besson, Chobert & Marie, 2011). Doelling & Poeppel (2015) report that entrainment to rhythm in the delta-theta range is enhanced in musicians, and that this entrainment was strongly correlated with years of musical training. Musacchia et al. (2007) show that music modulates the response to auditory stimuli at brainstem level, reporting, in musicians, an enhancement and earlier latency of the delta wave response (~8ms post-stimulus onset) to music and speech when compared to non-musicians. Once atypical delta brain responses have been associated to difficulties with phonology (e.g. Power et al., 2013, 2016), one could hypothesize that not only literacy but also music could change the speech auditory representations. In fact, a recent study using dyslexic musicians and non-musicians suggested that music training may help to overcome auditory sensory impairments (including rhythm perception) typically present in dyslexia (Bishop-Liebler et al., 2014). This study also reports that dyslexic musicians have a better overall performance in phonological awareness and word / pseudoword reading tasks than dyslexic non-musicians while they do not differ in other general intelligence tests (block design, matrices reasoning and similarities subtests from the Wechsler Adult Intelligence Scale). Taken together, this data may suggest that phonology enhancement is not confined to literacy acquisition or vocabulary development. Therefore, it is possible that other variables not assessed in our study (such as music training) can change the rhythmic structure of spontaneous speech. To explore this hypothesis, further research is needed.

Limitations and future research

The main limitations of the present study are the unbalanced ratio between men and women tested (total N = 46; 5 males) and the advanced age of the subjects what could possibly affect the generalizability of the results to other populations. Moreover, while the subjects in the highly literate group were recruited from a specific part of Lisbon, the literates with lower schooling as well as illiterates were recruited from different parts of the country. These regional differences (accent and manner of speaking) can possibly add variance to the produced speech properties not accountable to literacy. Also, the different socioeconomic status between highly literate subjects and low literates / illiterates makes it difficult to disentangle the effects of phonology development from other general language skills acquired in a more stimulating environment. Although our groups did not differ in age and did not report any hearing deficits, it would have been interesting to measure their hearing ability. In fact, the acoustic feedback is one of the external feedback loops that correct speech production features that show a deviation from the previous expectation (Hickok, 2012). Finally, although the S-AMPH model revealed to have sensitivity to the rhythm of linguistic units in this study, a thorough validation with a larger corpus of spontaneous speech would be important to confirm the rhythmic structure of the European Portuguese language. However, since speech rhythm may be cross-linguistically based on the grouping of stressed and unstressed syllables, it is expected that the phase relations between the linguistic AMs present in the S-AMPH provide useful data even for languages other than English (Leong, 2012; Goswami & Leong, 2013). Since in this work a link was established between phonologic skill and the physical properties of the speech signal, it could be interesting to study the temporal structure of speech in subjects with specific language impairments or dyslexia. A very recent study of mother's infant-directed speech to young infants showed that mothers, independently of being or not being dyslexic, do not hyperarticulate vowels when speaking to children at-risk for dyslexia, providing an altered early linguistic input that can affect the subsequent development of language skills (Kalashnikova, Goswami & Burnham, 2016). On the other hand, to further understand the links between neural processing and phase alignment in the speech signal, future research with functional imaging or electrophysiological techniques focusing on auditory processing areas (like the posterior superior temporal gyrus) could be enlightening.

Conclusions

Taken together, these findings support that literacy impacts significantly the temporal modulation structure of spontaneous adult-directed speech. Although the tendency for a literacy-induced energy shift from delta to theta AM did not reach significance, the phase alignment between slower and faster speech AMs was significantly enhanced by literacy and showed positive correlations with phonology, literacy and vocabulary measures. To our knowledge, this is the first study ever to use the modulation structure approach on illiterate speech and to show that literacy affects spoken language in terms of the physical/rhythmical properties of the output speech signal.

References

Anthony, J.L, Francis, D.J. (2015). Development of Phonological Awareness. *Current Directions in Psychological Science, Vol 14, Nº15,* 255-259.

Ardila, A., Bertolucci, P., Braga, L., Castro-Caldas, A., Judd, T., Kosmidis, M., Matute, E., Nitrini, R., Ostrosky-Solis, F., Rosselli, M. (2010). Illiteracy: The Neuropsychology of Cognition Without Reading. *Archives of Clinical Neuropsychology, 25,* 689-712.

Arnal, L, Poeppel, D., Hickok, G. *In* G. Hickok & L. Small (2015) *Neurobiology of Language,* Elsevier.

Atteveldt, N., Formisano, E., Goebel, R., Blomert, L. (2004). Integration of Letters and Speech Sounds in the Human Brain. *Neuron, Vol.43,* 271-278.

Bailey, P., Snowling, M. (2002). Auditory processing and the development of language and literacy. *British Medical Bulletin, 63,* 135-146.

Barreto, H, Moreira A.R., Ferreira, C. (2008). *Wechsler Adult Intelligence Scale (WAIS-III):* Versão Portuguesa. Lisboa, Portugal: Cegoc.

Besson, M., Chobert, J., Marie, C. (2011). Language and Music in the Musician Brain. *Language and Linguistics Compass, 5/9,* 617-634. DOI: 10.1111/j.1749-818x.2011.00302.x.

Blakemore, S-J., Frith, U. (2005). *The learning brain – lessons for education*. Blackwell Publishing.

Blau, V., Atteveldt, N., Ekkebus, M., Goebel, E., Blomert, L. (2009). Reduced Neural Integration of Letters and Speech Sounds Links Phonological and Reading Deficits in Adult Dyslexia. *Current Biology, 19,* 503-508.

Boemio, A., Fromm, S., Braun A., Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature Neuroscience, Vol.8, Nº3,* 389-395. DOI:10.1038/nn1409.

Brady, M., Armstrong, L., Mackenzie, C. (2006). An examination over time of language and discourse production abilities following right hemisphere brain damage. *Journal of Neurolinguistics, 19,* 291-310.

Brady, M., Mackenzie, C., Armstrong, L. (2003). Topic use following right hemisphere brain damage during three semi-structured conversational discourse samples, *Aphasiology*, 17:9, 881-904, DOI: 10.1080/02687030344000292.

Cameron, D., Grahn, J. (2014). Enhanced timing abilities in percussionists generalize to rhythms without a musical beat. *Frontiers in Human Neuroscience, Vol.8*. DOI: 10.3389/fnhum.2014.01003.

Canolty, R., Edwards, E., Dalal, S., Soltani, M., Nagarajan, S., Kirsch, H., Berger, M., Barbaro, N., Knight, R. (2006). High Gamma Power Is Phase-Locked to Theta Oscillations in Human Neocortex. *Science, 313,* 1626-1628. DOI: 10.1126/science.1128115

Carreiras, M., Seghier, M.L., Baquero, S., Estévez, A., Lozano, A., Devlin, J.T., Price, C.J. (2009). An anatomical signature for literacy. *Nature, Vol 461, 15,* 983-986. doi:10.1038/nature08461.

Castro, S.L, Caló, S., Gomes, I. (2007). PALPA-P: Provas de Avaliação da Linguagem e da Afasia em Português.

Castro-Caldas, A., Miranda, P.C, Carmo, I., Reis, A., Leote, F., Ribeiro, C., Ducla-Soares, E. (1999). Influence of learning to read and write on the morphology of the corpus callosum. *European Journal of Neurology, 6,* 23-28.

Castro-Caldas, A., Nunes, M.V, Maestu, F., Ortiz, T., Simoes, R., Fernandes, R., De la Guia, E., Garcia, E., Gonçalves, M. (2009). Learning orthography in adulthood: A magnetoencephalographic study. *Journal of Neuropsychology, 3,* 17-30.

Castro-Caldas, A., Petersson, K.M., Reis, A., Stone-Elander, S., Ingvar, M. (1998). The illiterate brain Learning to read and write during childhood influences the functional organization of the adult brain. *Brain, 121,* 1053-1063.

Castro-Caldas, A., Reis, A. (2000). Neurobiological Substrates of Illiteracy. *The Neuroscientist, 6,* 475-482.

Catani, M., Mesulam, M. (2008). The arcuate fasciculus and the disconnection theme in language and aphasia: History and current state. *Cortex, 44,* 953-961.

Chait, M., Greenberg, S., Arai, T., Simon, J., Poeppel, D. (2015). Multi-time resolution analysis of speech: evidence from psychophysics. *Frontiers in Neuroscience, Vol. 9,* doi: 10.3389/fnins.2015.00214.

Cohen, L., Dehaene, S. (2004). Specialization within the ventral stream: the case for the visual word form area. *NeuroImage, 22,* 466-476.

Crinion, J., Lambon-Ralph, M., Warburton, E., Howard, D., Wise, R. (2003). Temporal lobe regions engaged during normal speech comprehension. *Brain, 126,* 1193-1201, DOI: 10.1093/brain/awg104.

Damásio, A.R., Damásio, H. (1983). The anatomic basis of pure alexia. *Neurology, 33,* 1573-1583.

Damásio, H., Damásio, A.R. (1980). The Anatomical Basis of Conduction Aphasia. *Brain, 103,* 337-350.

Davidson, B., Worrall, L., Hickson, L. (2003). Identifying the communication activities of older people with aphasia: Evidence from naturalistic observation, *Aphasiology, 17:3*, 243-264, DOI: 10.1080/729255457.

De Cara, B., Goswami, U. (2003). Phonological neighbourhood density: effects in a rhyme awareness task in five-year-old children. *Journal of Child Language, Vol 30, 3,* 695-710.

Dehaene, S., Cohen, L., Morais, J., Kolinsky, R. (2015). Illiterate to literate: behavioural and cerebral changes induced by reading acquisition. *Nature Reviews, Vol.16,* 234-244.

Dehaene, S., Pegado, F., Braga, L.W., Ventura, P., Filho, G.N, Jobert, A., Daehaene-Lambertz, G., Kolinsky, R., Morais, J., Cohen, L. (2010). How Learning to Read Changes the Cortical Networks for Vision and Language. *Science, 330,* 1359-1365. DOI: 10.1126/science.1194140.

Ding, N., Patel, A., Chen, L., Butler, H., Luo, C., Poeppel, D. (2016). Temporal Modulations Reveal Distinct Rhythmic Properties of Speech and Music. *bioRxiv.* DOI: 10.1101/059683.

Ding, N., Simon, J. (2014). Cortical entrainment to continuous speech: functional roles and interpretations. *Frontiers in Human Neuroscience, Vol.8, 331,* 1-7.

Doelling, K., Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *PNAS.* DOI: 10.1073/pnas.1508431112.

Fisk, J., Sharp, C. (2004) Age-Related Impairment in Executive Functioning: Updating, Inhibition, Shifting, and Access, *Journal of Clinical and Experimental Neuropsychology*, *26:7*, 874-890.

Folstein, M., Folstein, S., & McHugh, P. R. (1975). Mini-Mental State: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*. *12*, 189-198.

Friedrici, A. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, Vol.6, Nº2,* 78-84.

Frith, U. (1998). Literally changing the brain. *Editorial – Oxford University Press,* 1011-1012.

Gaillard, R., Naccache, L., Pinel, P., Clémenceau, S., Volle, E., Hasboun, D., Dupont, S., Baulac, M., Dehaene, S., Adam, C., Cohen, L. (2006). Direct Intracranial, fMRI, and Lesion Evidence for the Causal Role of Left Inferotemporal Cortex in Reading. *Neuron, 50,* 191-204. DOI 10.1016/j.neuron.2006.03.031.

Ghitza, O, 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.

Giraud, A., Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience, Vol. 15, Nº4,* 511-518

Gireesh, E., Plenz, D. (2008). Neuronal avalanches organize as nested theta and beta/gamma-oscillations during development of cortical layer 2/3. *PNAS, Vol. 105, Nº21,* 7576-7581.

Gomes, I., Castro, S. L., & Lima, C. F. (2005) [on-line]. *Porlex, base lexical em Português Europeu (versão 2)*. Disponibilizado por Laboratório de Fala da FPCE Universidade do Porto em http://www.fpce.up.pt/labfala

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences, Vol. 15, Nº1,* 3-10.

Goswami, U. (2016). Educational neuroscience: neural structure-mapping and the promise of oscillations. *Current Opinion in Behavioural Sciences, 10,* 89-96. DOI: 10.1016/j.cobeha.2016.05.011.

Goswami, U., Barnes, L., Mead, N., Power, A., Leong, V. (2016). Prosodic Similarity Effects in Short-Term Memory in Developmental Dyslexia. *Dyslexia,* 1-18. DOI: 10.1002/dys.1535.

Goswami, U., Cumming, R., Chait, M., Huss, M., Mead, N., Wilson, A., Barnes, L., Fosker, T. (2016). Perception of Filtered Speech by Children with Developmental Dyslexia and Children with Specific Language Impairments. *Frontiers in Psychology, Vol. 7.* DOI: 10.3389/fpsyg.2016.00791.

Goswami, U., Leong, V. (2013). Speech rhythm and temporal structure: Converging perspectives? *Laboratory Phonology, 4(1),* 67-92. DOI: 10.1515/lp-2013-0004.

Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S., Scott, S. (2002). Amplitude envelope onsets and developmental dyslexia: A new hypothesis. *PNAS, vol.99, nº16,* 10911-10916.

Greenberg, S., Carvey, H., Hitchcock, L., Chang, S. (2003). Temporal properties of spontaneous speech – A syllable-centric perspective. *Journal of Phonetics,* DOI: 10.1016/j.wocn.2003.09.005

Guerreiro M, S. A., Botelho MA, Leitão O, Castro-caldas A, Garcia C. (1994). “Adaptação à população portuguesa da tradução do "mini mental state examination" (MMSE).” *Revista Portuguesa de Neurologia, 3(suppl 1)*: 9-10.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews, Vol 13,* 135-145.

Hickok, G. (2014). The architecture of speech production and the role of the phoneme in speech processing. *Language, Cognition and Neuroscience, 29:1*, 2-20, DOI: 10.1080/01690965.2013.834370.

Hickok, G., Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences, Vol.4, Nº4,* 131-138.

Hickok, G., Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition, 92,* 67-99.

Hickok, G., Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews, Vol.8,* 393-402.

Hogan, T., Bowles, R., Catts, H., Storkel, H. (2010). The influence of neighborhood density and word frequency on phoneme awareness in 2nd and 4th grades. *Journal of Communication Disorders, 44,* 49-58.

Indefrey, P., Levelt, W. (2004). The spatial and temporal signatures of word production components. *Cognition, 92,* 101-144. DOI:10.1016/j.cognition.2002.06.001.

Kalashnikova, M., Goswami, U., Burnham, D. (2016). Mothers speak differently to infants at-risk for dyslexia. *Developmental Science,* 1-15. DOI: 10.1111/desc.12487.

Kay ,J., Ruth ,L., Coltheart, M. (1996) Psycholinguistic assessments of language processing in aphasia (PALPA): An introduction, *Aphasiology*, 10:2, 159-180

Kim, J., Lee, S., Yoon, J. (2015). The distinctive effect of providing syllables in letter fluency. *Speech Communication, 70,* 42-48.

Kosmidis, M., Tsapkini, K., Folia, V. (2006). Lexical processing in illiteracy: effect of literacy or education? *Cortex, 42,* 1021-1027.

Lakatos, P., Shah, A., Knuth, K., Ulbert, I., Karmos, G., Schroeder, G. (2005). An Oscillatory Hierarchy Controlling Neuronal Excitability and Stimulus Processing in the Auditory Cortex. *Journal of Neurophisiology, 94,* 1904-1911.

Lawton, M.P., Brody, E.M. (1969). Assessment of Older People: Self-Maintaining and Instrumental Activities of Daily Living. *The Gerontologist, 9,* 179-186.

Lehongre, K., Ramus, F., Villiermet, N., Schwartz, D., Giraud, A. (2011). Altered Low-Gamma Sampling in Auditory Cortex Accounts for the Three Main Facets of Dyslexia. *Neuron, 72,* 1080-1090, DOI 10.1016/j.neuron.2011.11.002.

Leong, V. (2012). *Prosodic rhythm in the speech amplitude envelope: Amplitude modulation phase hierarchies (AMPHs) and AMPH models*. Doctoral dissertation, University of Cambridge.

Leong V, Kalashnikova M, Burnham D, Goswami U. (2014) Infant-directed speech enhances temporal rhythmic structure in the envelope. *InterSpeech*.

Leong V., Goswami U. (2015). Acoustic-Emergent Phonology in the Amplitude Envelope of Child-Directed Speech. *PLoS ONE, 10(12)*: e0144411. doi:10.1371/journal.pone.0144411.

Leong, V., Goswami, U. (2013). Assessment of rhythmic entrainment at multiple timescales in dyslexia: Evidence for disruption to syllable timing. *Hearing Research, http://dx.doi.org/10.1016/j.heares.2013.07.015.*

Leong, V., Goswami, U. (2013). Assessment of rhythmic entrainment at multiple timescalesin dyslexia: Evidence for disruption to syllable timing. *Hearing Research.* DOI: 10.1016/j.heares.2013.07.015.

Leong, V., Goswami, U. (2016). Difficulties in auditory organization as a cause of reading backwardness? An auditory neuroscience perspective. *Developmental Science.* DOI: 10.1111/desc.12457.

Levelt, W., Roelofs, A., Meyer, A. (1999). A theory of lexical access in speech production. *Behavioural Brain Sciences, 22:1,* 1-75.

Levelt, W., Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition, 50,* 239-269.

Liberman, I., Shankweiler, D., Fischer, F., Carter, B. (1973). Explicit Syllable and Phoneme Segmentation in the Young Child. *Journal of Experimental Child Psychology,* 1-24.

Liebler, P., Welch, G., Huss, M., Thomson, J., Goswami, U. (2014). Auditory Temporal Processing Skills in Musicians with Dyslexia. *Dyslexia, 20,* 261-279, DOI: 10.1002/dys.1479.

Lindfield, K., Wingfield, A., Goodglass, H. (1999). The Role of Prosody in the Mental Lexicon. *Brain and Language, 68,* 312-317.

Loureiro, C.S, Braga, L.W., Souza, L.N., Filho, G.N, Queiroz, E., Dellatolas, G. (2004). Degree of illiteracy and phonological and metaphonological skills in unschooled adults. *Brain and Language, 89,* 499-502.

Luo, H., Poeppel, D. (2007). Phase Patterns of Neuronal Responses Reliably Discriminate Speech in Human Auditory Cortex. *Neuron, 54,* 1001-1010.

Madureira, S., Moleiro, C., Verdelho, A., Guerreiro, M. (2008). *Escala de Atividades Instrumentais de Vida Diária (AIVD) Escala de Actividades Instrumentais de Vida Diária. Escala e Testes na Demência*. Lisboa: Novartis.

Marie, C., Magne, C., Besson, M. (2010). Musicians and the Metric Structure of Words. *Journal of Cognitive Neuroscience, 23:2,* 294-305. DOI: 10.1162/jocn.2010.21413.

Matute, E. (1986). Aphasia in illiterates. *Journal of Neurolinguistics, 2(1-2),* 115-130. DOI: 10.1016/S0911-6044(86)80007-0.

McCandliss, B.D., Cohen, L., Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences, Vol 7, Nº7,* 293-299.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition, 29,* 143-178.

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. *In* J. L. Metsala, & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Mahwah, NJ: Lawrence Erlbaum Associates.

Molinaro, N., Lizarazu, M., Lallier, M., Bourguignon, M., Carreiras, M. (2016). Out-of-Synchrony Speech Entrainment in Developmental Dyslexia. *Human Brain Mapping*. DOI: 10.1002/hbm.23206.

Morais, J., Bertelson, P., Cary, L., Alegria, J. (1986). Literacy training and speech segmentation. *Cognition, 24,* 45-64.

Morais, J., Cary, L., Alegria, J., Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition, 7,* 323-331.

Musacchia, G., Sams, M., Skoe, E., Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *PNAS, Vol. 104, Nº40,* 15894-15898.

Nittrouer, S., Studdert-Kennedy, M., McGowan, R. (1989). The emergence of phonetic segments: evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech and Hearing Research, Vol. 32,* 120-132.

Orange, J., Kertesz, A., Peacock, J. (1998). Pragmatics in frontal lobe dementia and primary progressive aphasia. *Journal of Neurolinguistics, Vol. 11, Nº1-2,* 153-177.

Pattamadillok, C., Knierim, I., Duncan, K., Devlin, J. (2010). How Does Learning to Read Affect Speech Perception? *The Journal of Neuroscience, 30(25),* 8435-8444. DOI:10.1523/JNEUROSCI.5791-09.2010.

Peter, V., Kalashnikova, M., Santos, A., Burnham, D. (2016). Mature neural responses to Infant-Directed Speech but not Adult-Directed Speech in Pre-Verbal Infants. *Scientific Reports, 6:34273,* 1-14. DOI: 10.1038/srep34273.

Peterson, G., Lehiste, I. (1960). Duration of Syllable Nuclei in English. *The Journal of Acoustical Society of America, Vol 32, Nº6,* 693-703.

Petersson, K.M, Reis, A., Askelof, S., Castro-Caldas, A., Ingvar, M. (2000). Language Processing Modulated by Literacy: A Network Analysis of Verbal Repetition in Literate and Illiterate Subjects. *Journal of Cognitive Neuroscience, 12:3,* 364-382.

Petersson, K.M., Reis, A., Castro-Caldas, A., Ingvar, M. (1999). Effective Auditory–Verbal Encoding Activates the Left Prefrontal and the Medial Temporal Lobes: A Generalization to Illiterate Subjects. *NeuroImage, 10,* 45-54.

Petersson, K.M., Silva, C., Castro-Caldas, A., Ingvar, M., Reis, A. (2007). Literacy: a cultural influence on functional left-right differences in the inferior parietal cortex. *European Journal of Neuroscience, Vol 26,* 791-799.

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as ‘asymmetric sampling in time’. *Speech Communication, 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.* 142-149. DOI:10.1016/j.conb.2014.07.005.

Power, A., Colling, L., Mead, N., Barnes, L., Goswami, U. (2016). Neural encoding of the speech envelope by children with developmental dyslexia. *Brain & Language,* 1-10. DOI: 10.1016/j.bandl.2016.06.006.

Power, A., Mead, N., Barnes, L., Goswami, U. (2012). Neural entrainment to rhythmically presented auditory, visual, and audio-visual speech in children. *Frontiers in Psychology, vol 3, 216,* doi: 10.3389/fpsyg.2012.00216

Power, A., Mead, N., Barnes, L., Goswami, U. (2013). Neural entrainment to rhythmic speech in children with developmental dyslexia. *Frontiers in Human Neuroscience, Vol.7, 777,* 1-19.

Qu, Q., Damian, M., Kazanina, N. (2012). Sound-sized segments are significant for Mandarin speakers. *PNAS, Vol. 109, Nº35,* 14265-14270. DOI: 10.1073/pnas.1200632109.

Ramus, F. (2002). Language discrimination by newborns Teasing apart phonotactic, rhythmic, and intonational cues. *Annual Review of Language Acquistion, 2,* 85-115.

Reis, A., Castro-Caldas, A. (1997). Illiteracy: A cause for biased cognitive development. *Journal of the International Neuropsychological Society, 3,* 444-450.

Sammler, D., Kotz, S., Eckstein, K., Ott, D., Friedrici, A. (2010). Prosody meets syntax: the role of the corpus callosum. *Brain, 133,* 2643-2655.

Schack, B., & Weiss, S. (2005). "Quantification of phase synchronization phenomena and their importance for verbal memory processes". *Biological Cybernetics 92*, 275-287.

Schlaug, G. (2001). The brain of musicians. A model for functional and structural adaptation. *Annals of the New York Academy of Sciences, 930,* 281-299.

Schlaug, G., Jancke, L., Huang, Y., Staiger, J., Steinmetz, H. (1995). Increased Corpus Callosum Size in Musicians. *Neuropsychologia, Vol. 33, Nº8,* 1047-1055.

Schmithorst, V., Wilke, M. (2002). Differences in white matter architecture between musicians and non-musicians: a diffusion tensor imaging study. *Neuroscience Letters, 321,* 57-60.

Schotten, M.T., Cohen, L., Amemiya, E., Braga, L.W., Dehaene, S. (2014). Learning to Read Improves the Structure of the Arcuate Fasciculus. *Cerebral Cortex, 24,* 989-995. doi:10.1093/cercor/bhs383.

Schroeder, C., Lakatos, P. (2008). Low-frequency neuronal oscillations as instruments of sensory selection. *Cell Press.* DOI:10.1016/j.tins.2008.09.012.

Seidenberg, M., Tanenhaus, M. (1979). Orthographic Effects on Rhyme Monitoring. *Journal of Experimental Psychology: Human Learning and Memory. Vol.5, Nº6,* 546-554.

Serniclaes, W., Ventura, P., Morais, J., Kolinsky, R. (2005). Categorical perception of speech sounds in illiterate adults. *Cognition, 98,* B35-B44.

Shannon, R., Zeng, F., Kamath, V., Wygonsky, J., Ekelid, M. (1995). Speech Recognition with Primarily Temporal Cues. *Science, Vol. 270,* 303-304.

Sieben, K., Hartung H., Wolff, A., Hanganu-Opatz, I. (2013) The melody of the immature brain. *e-Neuroforum, 4,* 11-17. DOI: 10.1007/s13295-013-0037-1.

Trainor, L., Shahin, A., Roberts, L. (2009). Understanding the Benefits of Musical Training Effects on Oscillatory Brain Activity. *Annals of the New York Academy of Sciences, 1169,* 133-142. DOI: 10.1111/j.1749-6632.2009.04589.x.

van der Mark, S., Bucher, K., Maurer, U., Schulz, E., Brem, S., Buckelmuller, J., Kronbichler, M., Loenneker, T., Klaver, P., Martin, E., Brandeis D. (2009). Children with dyslexia lack multiple specializations along the visual word-form (VWF) system. *NeuroImage, 47,* 1940-1949.

Webster, P., Plante, A. (1995). Productive phonology and phonological awareness in preschool children. *Applied Psycholinguistics, 16,* 43-57.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale—3rd Edition (WAIS-3®).* San Antonio, TX: Harcourt Assessment.

Wernicke C. (1874) Der aphasische Symptomenkomplex. *In* Eggert GH (1977),. *Wernicke's Works on Aphasia*. The Hague: Mouton.

Zhang, R. (2012). Rhetorical Analyses of English Proverbs. *US-China Foreign Language, Vol. 10, Nº 3*, 1005-1009.

Ziegler, J.C., Goswami, U. (2005). Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory. *Psychological Bulletin, Vol. 131, Nº1*, 3-29.