U.PORTO
FACULDADE DE PSICOLOGIA
E DE CIÊNCIAS DA EDUCAÇÃO
UNIVERSIDADE DO PORTO

Paulo Branco

Musical expertise and brain plasticity
an ERP study on speech parsing

Master Thesis
Psychology
2012

University of Porto

Faculty of Psychology and Education Sciences

MUSICAL EXPERTISE AND BRAIN PLASTICITY
AN ERP STUDY ON SPEECH PARSING

Paulo Branco

October 2012

Thesis supervised by Professor São Luis Castro and presented at the University of Porto, Faculty of Psychology and Education Sciences for the Master’s degree in Psychology

ACKNOWLEDGMENTS

This work is the result of a hard but rewarding experience. Several persons were involved in the elaboration of this thesis, and as such I must share some words of acknowledgment. Professor São Luis Castro, my supervisor, was essential for the development of all this work. Her knowledge aided me in many ways, and the amount of things I learned is too much to put into words. I am grateful for all the support, the constructive criticism, the hours spent reviewing my work and for believing in me since the start. Thanks to her, I was able to grow as a researcher, and I was given all the tools to do so. I am grateful for the trust and commitment into this thesis, and I hope this is the start of a new stage in my personal development in science. Dr. Susana Silva was also important for my growth as a researcher. Her enthusiasm and honesty were truly an advantage for someone as naive as I was (am) back then at the start of this project. Enormous thanks to Professor Fernando Barbosa and Professor João Marques-Teixeira, for being so receptive about this project and for allowing me into the neuropsychophysiology laboratory to do this research. I was blessed to be able to work in a close collaboration between two laboratories, and the exchange of knowledge between methodologies was very fulfilling. Thanks to Dr. Mireille Besson, Dr. Clement François, Dr. Julie Chobert, Dr. Radouane El Yagoubi and Prof. Majid Himmi for giving me important feedback about this work. It was a great experience to be able to discuss it with such knowledgeable researchers in this field! Professor Selene Vicente was also a source of inspiration and her classes and research activities motivated me to learn more. Her commitment to the students and research is an example to follow. Thank you very much! I must also thank the Speech Lab colleagues, for all the constructive feedback and for receiving me so well. To name a few, special thanks to Dr. César Lima, Dr. Rui Alves, Marisa Filipe and Manuela Cameirão, who always helped when I needed. Thanks to Dr. Pedro Almeida, Dr. Fernando Santos, Joana Vieira and Tiago Paiva from neuropsychophysiology laboratory for all the constructive feedback, and for teaching me a lot on EEG methods.

Finally, a word of appreciation to all my friends and girlfriend who for countless hours listened to my difficulties and expectations, but always supported me to follow my dream to become a researcher. Thank you for your patience and support!

"...a mocidade que vai passando, recheada
de entusiasmo e sonho, nem por sombras
faz ideia da amargura que a espera.
Felizmente para estes doutores o
espectáculo do envelhecimento não será
tão pungente como o meu.
A maior parte deles não ficará numa
cidade de jovens a vê-los sempre iguais,
com olhos sempre diferentes."

Miguel Torga, Diário V

ABSTRACT

Introduction. Expertise in the domain of music can have several nontrivial consequences. According to the OPERA hypothesis (Patel, 2011), neural plasticity is likely to occur when musicians use shared neural resources in music and language. Our question is whether musical expertise affects phrasing processes in speech. To this end, we examine Closure Positive Shift, CPS, an electrophysiological component in event-related potentials that reflects processing of prosodic boundaries, in musicians and non-musicians.

Method. Sixteen musicians (at least 7 years of musical training; 9 women, mean age 20.2 yrs) and 16 non-musicians (10 women, mean age 20 yrs) listened to short sentences and were asked to identify whether a probe word was presented. The sentences contained one or two phrase boundaries, and were spoken or hummed.

Results. The ERP traces differed between phrasing conditions, with evidence of CPS over one boundary contrasting with no CPS in sentences without the boundary. In musicians compared to non-musicians, the CPS latency was faster in spoken sentences and CPS amplitude was larger in hummed sentences. Hummed sentences elicited a more frontal distribution than spoken ones in both groups, with left lateralized topography in musicians.

Discussion. Taken together, these findings indicate that musical expertise may facilitate phrasing processes in speech. The lateralization effect fits in well with fMRI evidence showing that left brain areas are more activated when processing boundaries in hummed sentences.

Keywords: musical expertise, neural plasticity, closure positive shift, phrasing

RESUMO

Introdução. A expertise no domínio da música pode ter consequências a nível da linguagem. De acordo com a hipótese OPERA (Patel, 2011), ocorre plasticidade neuronal quando músicos usam recursos neuronais partilhados entre a música e a linguagem. A nossa questão é de saber se a expertise musical afecta o tratamento prosódico, concretamente a segmentação implícita de fronteiras prosódicas. Com este objectivo, examinamos o Closure Positive Shift, ou CPS, uma componente electrofisiológica nos potenciais relacionados com eventos que reflecte o processamento de fraseamento prosódico, em músicos e não-músicos.

Método. Dezasseis músicos (com pelo menos 7 anos de formação musical; 9 mulheres, idade média 20.2 anos) e 16 não-músicos (10 mulheres, idade média 20 anos) tinham como tarefa identificar em frases curtas se uma palavra-alvo tinha sido apresentada por via auditiva. As frases continham uma ou duas fronteiras prosódicas, e eram faladas ou ditas em boca fechada (bocca chiusa).

Resultados. Os traçados ERP foram diferentes nos dois tipos de frase, evidenciando um CPS contrastando com a sua ausência conforme estivesse ou não presente a fronteira prosódica. Para os músicos em comparação com os não-músicos, o tempo de latência do CPS foi mais rápido nas frases faladas, e amplitude foi maior nas frases em boca fechada. Para os dois grupos, as frases em boca fechada tiveram uma distribuição topográfica mais frontal do que as faladas, e lateralizada à esquerdo nos músicos.

Discussão. Estes resultados indicam que a experiência musical pode facilitar os processos de fraseamento na fala. O efeito de lateralização é concordante com resultados de ressonância magnética funcional mostrando que há maior activação em regiões do hemisfério esquerdo no tratamento prosódico de frases deslexicalizadas.

Palavras-chave: expertise musical, plasticidade neuronal, closure positive shift, tratamento prosódico

RESUMÉ

Introduction. L'expertise dans le domaine de la musique peut avoir des conséquences hors la musique. Selon l'hypothèse OPERA (Patel, 2011), la plasticité neuronale est susceptible de se produire lorsque les musiciens utilisent des ressources neuronales communes à la musique et au langage. Notre question est de savoir si l'expertise musicale affecte les processus de traitement prosodique, notanment la segmentation implicite de groupes intonationnels. Nous examinons Closure Positive Shift, CPS, une composante électrophysiologique des potentiels évoqués qui reflète le traitement des frontières prosodiques, chez les musiciens et non-musiciens.

Méthode. Seize musiciens (au moins 7 ans de formation musicale; 9 femmes, âge moyen 20.2 ans) et 16 non-musiciens (10 femmes, âge moyen 20 ans) ont écouté des phrases courtes et ont été invités à identifier si un mot sonde a été présenté. Les phrases contenaient un ou deux groupes intonationnels, et étaient parlées ou fredonnées.

Résultats. Les traces ERP diffèrent entre les conditions de groupe intonationnel, en montrant un CPS sur une frontière intonationnelle qui contraste avec l'absence de CPS sue les phrases sans frontière. Pour les musiciens en comparaison aux non-musiciens, le temps de latence de CPS a été plus rapide dans les phrases parlées et l' amplitude plus grande dans les phrases fredonnées. Pour les deux groupes, les phrases fredonnées ont eu une topographie plus frontale que les parlées, que était latéralisée à gauche chez les musiciens.

Discussion. Pris ensemble, ces résultats indiquent que l'expertise musicale peut faciliter les processus de segmentation implicite de groupes intonationnels. L'effet de la latéralisation correspond bien à de données en IRMf montrant que les zones du cerveau gauche sont plus activées lors du traitement prosodique des phrases fredonnées.

Mots-clés: expertise musicale, plasticité neuronale, closure positive shift, traitement prosodique

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ANOVA	Analysis of variance
CPS	Closure Positive Shift
CPSm	Closure Positive Shift Music
dB	Decibel
EEG	Electroencephalography
e.g.	Exempli gratia, for example
ERP	Event-related potential
et al.	Et alii, and others
F_0	Fundamental frequency
fMRI	Functional Magnetic Resonance Imaging
Hz	Hertz
i.e.	Id est, that is
IPh	Intonational Phrase
IPB	Intonational phrase boundary
MEG	Magnetoencephalography
mm	millimeters
ms	Milliseconds
ns	Non-significant
ROI	Regions of interest
RT	Reaction time
s	Seconds
SD	Standard deviation
µV	Microvolts

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

1. Introduction

"Words give rise to the music, and music develops and reinforces the language."

Richard Wagner (1813-1883)

Language plays a central role in human communication. It is one of the few tools that differentiates us from other species by allowing to think in objective terms, express ideas and feelings or to better understand our world (Patel, 2008). Music also allows us to communicate and has maybe other functions such as expressing emotions. Although superficially both seem quite different, they share a few interesting characteristics. Music is an important aspect of human evolution, as has been shown in phylogenetic studies. As early as the XVIII century, Rousseau (1781/1981) believed that music was the origin of language, where language would be the rationalization and organization of music. Darwin (1871/1974) suggested that music and language have similar origins, although according to him it was language that developed into music. Music has been considered relevant for human evolution by allowing individuals to communicate critical information to survive, such as localization, and by allowing reproductive calls to the opposite sex. Such relationships between language and music, even if questionable by today's scientific standards, suggest how close in purpose they may be. In an extensive review of language and music similarities, Patel (2008) suggests a division into sub-categories. In language, these categories are phonetics and phonology, including prosody, syntax and semantics. In music, they are pitch, rhythm, melody, syntax and meaning. Although we can find many similarities and differences between each of these, the main focus of this work is on some of the common acoustic elements are used to create perceptual groupings, or phrases.

If we divide music and language into their constituents, various parallels can be identified at the acoustic level. In this work, we focus on pitch. Frequency, the acoustic correlate of pitch, refers to the number of cycles per second in a sound and is measured in Hertz. As a result of resonation, there are no pure sounds in our universe. Most sounds are composed by harmonic multiples. The multiples can be filtered down to the fundamental frequency, F_0, the lowest frequency on top of which the other harmonics superimpose. In speech, F_0 is generated by the vibration of the vocal cords, higher frequency eliciting higher pitch perception. After interacting with the physical barriers of the glottis, the larynx and velum, and articulators such as tongue and lips, harmonics are created. The requirements to produce a tone in speech are to have a resonating chamber (pharyngeal, nasal and oral cavities), an energy source (the air flow), a vibrating mechanism (the vocal cords), and a propagation mechanism (the mouth); these provide aerodynamic conditions to create sound (Catford, 2001). A musical instrument behaves in an analogous way. A violin, for example, reacts to the vibration of the cords in contact with the bow. The sound goes into the amplification chamber (the wooden compartment) and is converted into the desired tone. As with the vocal cords, the faster a string vibrates, the higher the perceived pitch.

Pitch interacts with other variables to create prosodic phrasing: it is the grouping of various pitch points over a certain time that make up a "comprehensible" acoustic chain. The other variables are duration, rhythm and timing. Duration is the time during which there is vibration of a sound source, usually measured in milliseconds. Rhythm is the subdivision of each sound sequence in a regular or irregular way. Timing refers to the acoustic landmarks that segment a temporal sentence in language and in music (Kraus & Chandrasekaran, 2010). In language, timing is dependent on articulator movements (e.g. tongue, lips) and allows to control the formants, while in music it is related to the duration of sounds and their perceptual manipulation into rhythm. These are important to convey suprasegmental representations of spoken sentences and musical phrases, including those related to emotional prosody. For example, in music and language excitement is conveyed by fast and sharp acoustic cues (Besson & Friederici, 1998; for a review on other emotions see Banse & Scherer, 1996).

Another important shared acoustical attribute is intensity, defined by the amplitude of the sound wave and measured in decibels (dB). The higher the amplitude, the louder a stimulus sounds. This is an important feature for stress patterns both in music and language. Finally, the timbre, commonly known as sound color, allows us to differentiate between two sounds with the same pitch. In language as in music it is dependent on physical attributes of the voice or instrument, such as the amplifying mechanisms, the articulators, and the harmonics that arise from the resonation within the system. This allows us to identify someone's characteristic voice, as well as to differentiate between a piano and a guitar when both play the same tonal frequency. It is also an important cue to distinguish phonemes due to the interaction of resonance characteristics with harmonic composition, which amplify certain frequencies and attenuate others, building onto speech formants (Kraus & Chandrasekaran, 2010). From the cursory review of those shared acoustic attributes, we can apprehend that although music and language are obviously different in some respects, the “matter” used to create music and speech is very similar. We may thus ask the question: if, like Wagner suggested, words give rise to music and music develops and enhances language, how would a music expert process language?

Musical expertise

The definition of a musician as a subject who gained expertise on musical performance is rather ambiguous – what exactly is a musician? A self-taught multi-instrumentalist with no knowledge on musical theory, an orchestra soloist, or both? Can a person who listens to music everyday be called an expert in “listening and perceiving music”? Where exactly do we draw the line on what is musical training? Such questions, although hard to address, are central to understand how music is learnt. A good starting point is to look at some of the definitions of expertise.

A key factor to develop expertise in any domain is deliberate practice (Ericsson, Nandagopal & Roring, 2005). However, other factors such as general intelligence and domain-specific skills have been referred as important factors in musical training (Detterman & Ruthsatz, 1999). Ruthsatz, Detterman, Griscom and Cirullo (2008) have shown that the three variables, general intelligence, musical skills and deliberate practice, affect musical achievement but only on starting musicians. The performance of professional orchestra musicians is strongly related only to deliberate practice. In sum, to be an expert in music one must have an extended deliberate practice, though general intelligence and ability also play a role. A reasonable view is then to define a music expert as someone who has a musical degree or has extensive formal training in music.

Having specified what can be considered a music expert, what is different in experts? In a review of expertise, Chi, Glaser and Farr (1988) suggested that experts have specific cognitive processing advantages over novices, such as perceiving large meaningful patterns, being faster in specific tasks while showing less errors, having superior short and long term memory in the domain, representing a certain problem at a deeper level than novices, spending more time analyzing problems qualitatively and having stronger self-monitoring skills. Cellier, Eyrolle and Marine (1997) proposed that experts also have greater skill in anticipating processes, predicting them by resorting to domain-specific cues (for a review of other characteristics, see Farrington-Darby & Wilson, 2006). So of course musicians have advantages in music processing. Could these advantages be generalized outside of the music domain?

Brain plasticity, musical expertise

Brain plasticity is well-known to occur in musicians (e.g. Gaser & Schlaug, 2003a, Kraus & Chandrasekaran, 2010, Schlaug, 2001). The early age at which musical training usually starts is in line with critical periods of brain development, that is likely to adapt to these challenges (Penhune, 2011; Kraus & Chandrasekaran, 2010). For example, studies on brain plasticity show that with just a few weeks of focused practice, cortical areas corresponding to finger movement are increased (Pascual-Leone, Grafman & Hallett, 1994). Karni et al. (1995) further uncovered that musical training elicits morphological changes in the primary motor cortex. Also, other structures appear larger in volume in musicians, such as the corpus callosum and the cerebellum (Schlaug, 2001). Diffusion tensor imaging analysis shows that the neural connections in the internal capsule appear in greater numbers in musicians than in non-musicians (Han et al, 2009). On more cognitive-related structures, the planum temporale is more asymmetric in musicians than non-musicians, but only in musicians with absolute pitch (Schlaug, 2001). Furthermore, musicians have more gray matter volume in the sensory-motor cortex (Gaser & Schlaug, 2003b, Schlaug, Norton, Overy & Winner, 2005), auditory and visual regions (Gaser & Schlaug, 2003a), an increase which is correlated with years of musical experience. These anatomical differences clearly show that musicians have fine-tuned neural mechanisms associated with their expertise. In addition to these findings, the correlation between the onset of musical training and structure volume is evidence that such effects cannot be fully explained by innate differences in subjects who follow musical studies.

At the functional level, studies show the impact of musical training in brain activations. Various fMRI studies have shown that musicians, compared to non-musicians, had increased activations in areas such as Heschl gyrus (Schneider et al, 2002), the planum temporale bilaterally (Ohnishi et al., 2001), the frontal operculum, the right inferior frontal cortex, the anterior part of the superior temporal gyrus (Koelsch, Fritz, Schulze, Alsop & Schlaug, 2005) and even the hypoccampus (Herdener et al., 2010). Thus, musical training seems to influence not only the morphology of several brain areas, but also their functional role in various tasks.

Musical expertise and language

So far we have seen what musical expertise is and how it may affect the brain in both anatomy and function (e.g. Gaser, 2003a, Schlaug, 2001). Also, we have seen how some of the areas associated with musical expertise are also involved in language processing (e.g. Knosche et al., 2005, Marques, Moreno, Castro, & Besson, 2007, Moreno et al., 2009, Schon, Magne, & Besson, 2004). A growing number of empirical studies have aimed to specify how musically trained brains may process information in a specialized manner. These can be grouped into studies on higher level cognitive functions at the cortical level, and studies on low-level automated processes at the subcortical level.

Studies on higher level functions have analyzed how pitch detection and discrimination may be enhanced by musical training and musical expertise. In ERPs, using a pitch incongruity detection paradigm, Marques et al. (2007) showed that musicians were on average 300 ms faster to detect incongruous endings where pitch was manipulated increasing F_0. If the incongruity was small (35% increase) musicians showed a larger detection rate. These results are in line with findings by Schon et al. (2004), who showed that musicians process pitch more efficiently in language and music. Several other studies (c.f. Magne et al., 2007; Marie, Magne, & Besson, 2011, Moreno & Besson, 2006) have added further evidence on how pitch processing is enhanced by musical expertise, by improving pitch detection and discrimination and by reducing error rates in various tasks. However, whether musicians were trained to process pitch information more efficiently, or whether those were innate abilities that enabled them to become musical experts, is debatable. To explore this, Moreno et al. (2009) performed a randomized, controlled trial with eight-year-old children, half receiving musical training and half receiving painting training for 6 months. The children had training sessions of 75 minutes, twice a week. They were tested at the start of the program and after 6 months of training with neurocognitive tests and with event-related potential methods. In an incongruity detection taks, the children with music training had a larger N300 ERP after the program, while the painting group did not. As for the behavioral measures, the detection rate was also improved in the children with musical training on small pitch variations. Furthermore, reading was also significantly improved in the musical group. These results showed that with just 6 months of training, both the behavioral and ERP results were significantly different between painting and musical training groups and the respective baseline measures. From these studies, we can conclude that musical training changes how pitch is processed, an acoustic parameter that is shared between music and language.

Temporal information is also more efficiently processed by musicians. Marie et al. (2011) analyzed ERP responses to the manipulation of metric and semantic information in a sentence, and showed that musical expertise influences the different stages of linguistic metrical processing. They found an increase in the automatic detection of the syllable (larger P200 amplitude), an integration of metric information (larger N400 amplitude), and a more efficient analysis of metric violations (P600 and late positivity effects). Magne et al. (2007) also showed that metric information is influenced by musical expertise. When sentences were artificially manipulated to create unexpected syllable lengthening that would disrupt the metrical structure while preserving pitch and timbre, musicians showed substantially less errors in the experimental task than non-musicians. Also, the P200 component (perceptual, low-level processing) in response to syllable lengthening was significantly larger in musicians. Put together, these results suggest that musical training influences selected aspects of speech perception.

Research in subcortical functions, namely, the *frequency-following responses* originated in the brainstem's nuclei, show that this wave pattern aligns temporally with acoustical information. This synchrony goes as far as to represent the three main domains shared by music and language: timing, pitch and timbre (Kraus, Skoe, Parbery-Clark & Ashley, 2009). Such findings are related to the codification of low-level information, because they analyze the acoustical information received in the cochlea and convert it into higher-level information, projecting it to cortical regions for further processing. Surprisingly, musical training seems to aid in this process. The brainstem of musicians seems to synchronize more efficiently with pitch and timing than that of non-musicians (Kraus & Chandrasekaran, 2010; Wong, Skoe, Russo, Dees, & Kraus, 2007). Such findings show that the differences between musicians and non-musicians may occur not only in higher-level processes but also on faster and automated processes that subtend acoustic processing in music and speech.

The OPERA hypothesis

Kraus and Chandrasekaran (2010) argued that since music shares many resources with language, most notably pitch and timing perception, the amount of time used in processing these cues could improve the shared neuronal pathways. This claim has been made based on a series of studies that evidenced how brainstem responses to pitch information are more efficient in musicians than non-musicians, and how such results strongly correlate to the years of musical training (i.e., 7 years of training have better results than just 2 years of training). Building on such empirical evidence, Patel suggested a comprehensive model of speech perception and how musical training might influence this system. Such proposal is termed the OPERA hypothesis (Patel, 2011, 2012). This hypothesis argues that in order for musical training to influence speech processing, specifically, the neural encoding of speech, it has to fulfill five conditions: overlay, precision, emotion, repetition and attention.

The overlay of neural circuits between music and language is critical. The auditory pathways are cortical and subcortical, with ascending as well descending connections between the cochleae nuclei, the superior olivary nuclei, the inferior coliculus, and the medial genicular body, projecting to the primary auditory cortex bilaterally (Patel & Iversen, 2007). As Kraus and Chandrasekaran (2010) show, these subcortical regions are critical for low-level decoding of perceptual attributes of both speech and music (i.e. periodicity, as the acoustical correlate of pitch). If at a cortical level these pathways may diverge and contribute to high levels of specificity, at the subcortical level, due to the restrictions in the number of neurons, areas and connections, the pathways for music and speech acoustical attributes are shared (Patel, 2011).

Precision is the second condition: in order for brain plasticity to occur, music must place higher demands on the auditory system than speech would. In language, studies show that even though pitch has many linguistic functions (i.e. boundary processing, emphasis and focus) it does not play an essential role in speech comprehension as subjects use cues such as the context or phonetic information to parse the speech chain (Patel, 2011). In contrast, musicians must keep track of pitch at all times. A mere 6% deviation from a certain tone results in a semitone, which is the immediate anterior note on a Western musical scale. This plays a special role in instruments where there is no physical boundary of tones, such as the violin or the cello, where the sense of tone is crucial for precise playing. Interestingly, it is the deficit in pitch detection that is presumably one of the underlying causes of amusia (Peretz, Cummings, & Dube, 2007). Another attribute crucial to music is timing. Timing is related to the perception of how correct, on a given time scale, a tone is. This gives temporal "meaning" to a musical phrase. Timing is a crucial aspect for musical performance and it is highly developed in musicians. Studies show that, when timing is manipulated to convey emotional meaning to a musical phrase, limbic and paralimbic areas related to emotional processing are activated, when compared to mechanical (deadpan) sentences (Chapin, Jantzen, Kelso, Steinberg & Large, 2010). These examples show that music training develops precision in pitch and timing in music. Since these are processed by shared neural circuitry, musical training would also enhance the precision of pitch encoding in speech (Patel, 2012).

The emotion category is the ability to be successfully engaged into learning music (and not by executing emotional pieces per se). According to the OPERA model, to engage in motivated and pleasurable musical practice is a necessary condition. The motivation to learn music by positive reinforcement, but also by being able to communicate through music, can be crucial for engaged learning, thus satisfying the requirement for deliberate practice, that involves repetition and attention. These two final categories are central to define expertise. It is possible to not have formal music instruction and still be involved in informal learning by passively listening to music. However, for brain plasticity to occur (and therefore, for music training to aid in speech encoding), there has to be focused attention into learning and perceiving the acoustic attributes found in music. Studies by Fritz, Elhilali and Shamma (2005) and Polley, Steinberg and Merzenich (2006) on animal models show that actively attended stimuli facilitate brain plasticity. Focused attention and repetition are therefore key elements to guide brain plasticity mechanisms by enabling musicians to improve selective competences in a way that transfers to language. In sum, it is only when subjects perform a task attentively, deliberately, and repeatedly, with a learning curve that is highly demanding, that induced brain plasticity may simultaneous affect language and music.

OPERA hypothesis and phrasing

Musical training influences various levels of information processing, and the OPERA hypothesis argues that if the conditions are met, then those skills transfer to language. In this study, we are interested in how musical expertise may affect the grouping and segmenting of sentences. To do so, we must discuss which acoustic cues are used to segment sentences, and how, accordingly to this hypothesis, this may be affected by musical training.

A sentence is usually built by intonational phrases (IPh). An IPh is a segment of speech with a single pitch and specific rhythm and has a few characteristic acoustic properties: it must have at least one boundary tone at the right edge, a pause that separates one IPh from the next, an F_0 descending pattern towards the end of the sentence and also lengthening of the final syllable. It is also affected by syntax and semantics (Wang & Hirschberg, 1992). Schonberg (1967) as cited in Silva (2005), also argues that perception of phrase ending in musical sentences depends on pauses, melodic relaxation by the drop of pitch, availability of less notes, rhythm reduction and shorter intervals. Patel, Peretz, Tramo and Labrecque (1998) also illustrate how musical phrases have a characteristic lengthening of the last notes. From these characterizations, we can observe how both acoustic cues that guide the production and perception (and furthermore segmentation) of intonational phrases are similar to those found to be more efficiently processed by musicians. Therefore, accordingly to OPERA, musical expertise would benefit the process of grouping and segmenting sentences by the mechanisms of brain plasticity, if the five conditions are met. We have seen previously how emotion, repetition and attention are critically used by musicians during their training. For this argument, we will discuss the two first conditions of OPERA hypothesis applied to phrasing.

The first condition, overlay, requires that phrasing in music depends on the same neural circuitry as speech. As seen previously, pitch and timing are encoded in the brainstem through the auditory neural pathways (see Kraus & Chandrasekaran, 2010 for a detailed explanation), and both are nuclear to perceive an intonational phrase accordingly to linguistic theories (Wang & Hirschberg, 1992). There is some overlap in the circuitry responsible for conversion of low-level acoustic information into higher level, cortically integrated percepts. Also, studies have shown that several brain areas are shared in music and language processing. The Broca region was also found to be of utmost importance in processing musical stimuli (Vuust et al, 2006; for a review of other areas, see Abrams et al., 2011). Furthermore, if brain areas are common, then it is safe to assume that impairments in prosody should also reflect in music processing. In fact, Patel et al. (1998) show that brain damaged patients in analogous areas had difficulties processing both music and language, with problems ranging from the detection of pauses (boundaries) to the discrimination of musical patterns. Thus there may be an overlap of areas responsible for some processes of both language and music. Therefore, the first condition is met.

The second condition, precision, is that musicians use the same processes as non-musicians but at more demanding level. Phrasing is of utmost importance in music. As Silva (2005) points out, a good musical performer must excel in phrasing, that is, to execute musical phrases based on implicit (yet correct) knowledge on a set of musical sentences in a specific musical piece. In fact, "a normal musician analyses the score, searches for cadence and then phrases accordingly to this; a talented musicians 'sees' the phrase immediately" (V.A, as cited in Silva, 2005, pp. 48). The perception and anticipation of a boundary in a musical phrase is therefore critical to successful performance in music. Furthermore, music has to maintain a constant *tempo*, placing higher demands on the mechanisms that allow musicians to "read" a boundary and act accordingly, not only by the use of informational cues (like the musical sheet) but also by intuition.

In sum, we can see how, accordingly to this theory, musical expertise is expected to promote brain plasticity and consequently, improve phrasing processing. We have also seen how music and language share some of the important acoustic attributes that allow subjects to successfully parse an intonational phrase, segment it, and then move attention to the next chunk of information. It is then pertinent to clarify whether such an hypothesis would be confirmed in experiments about phrasing processes in musicians and non-musicians.

Phrasing in the brain

fMRI and EEG techniques are used to study the neural mechanisms of speech production and perception. Also, both have been employed with success to study phrasing processes. Each has several known limitations and advantages. Although EEG has a good temporal resolution, it is difficult to trace the ERP to its origins and therefore to understand exactly which parts of the brain are recruited (Luck, 2005). With fMRI, we can indirectly measure brain activation, but only in large time-windows, losing temporal criteria. The study of brain areas responsible for the segmentation of sentences using fMRI is still underdeveloped. However, a study by Ischebeck et al. (2008) tried to disclose what brain regions would be critical for phrasing, in spoken and hummed sentences. The experiment found similar areas in both conditions: posterior rolandic operculum, supra-temporal and Heschl's gyrus, bilaterally. However, when hearing hummed sentences, the phrasing would appear lateralized to the left hemisphere. This effect was interpreted as an attention focus on segmental information, required in the experiment.

In contrast, the study of phrasing with EEG methods has proliferated over the last decade. Using ERP, a lot of information has been added to our knowledge on the domain (cf. Knosche et al., 2005, Steinhauer et al., 1999). The present study focuses on a specific component – the closure positive shift - that reflects the processing of phrasing in both speech and music.

Closure Positive Shift

The Closure Positive Shift, CPS, is an electrophysiological component in ERPs found by Steinhauer et al. (1999) that reflects processing of prosodic boundaries in speech. It has a central topography and appears bilaterally at around 300-700 ms after the end of an intonational phrase (Steinhauer, 2003). It does not depend on other well-established auditory ERP components, such as the P600 or P800, or even exogenous components such as P50 and P200 (Mannel & Friederici, 2011, Steinhauer, 2003). This component does not rely on lexical information, as stimuli with prosodic-only information (by the expression of mean F_0, using a computerized filter) elicited a similar response as for speech (Steinhauer & Friederici, 2001; Steinhauer, 2003). Even sentences without pauses elicit a CPS, thus showing that this component is a response to the end of an intonational phrase, probably activating phonological representations of the sentence based on syllable lengthening and pitch modulation (Steinhauer, 2003). There are three main proposals on the functional role of CPS. Steinhauer et al. (1999) suggest that the CPS is a structural marker of the end of an intonational phrase. Knosche et al. (2005) argued that it might be related to attentive processes as a way to integrate information between different sentences. Kerkhofs, Vonk, Schriefers and Chwilla (2007) propose that this component reflects a brain mechanism to organize sentence information.

Several experiments have revealed that the topography of CPS depends on the type of stimuli. Pannekamp, Toepel, Alter, Hahne, and Friederici (2005) used normal speech, jabberwocky speech (nonwords in correct syntactic sequences), pseudo-sentences (no syntactic agreement, but phonologically correct pseudo-words) and hummed speech (no lexical cues). The results show a correspondence between the amount of lexical information and topography: the less lexical information, the more frontal the component. Jabberwocky sentences were less frontal that pseudo-sentences and both were less frontal than hummed sentences. These results are suggestive of a complex neural network in which the brain requires more frontal resources, typically used for long-term planning and executive control, to process less meaningful speech (Pannekamp et al., 2005).

CPS is also elicited by silent reading. Steinhauer (2003) showed that commas can have the same effect as an acoustic pause. In a task where subjects were asked to silently read a set of sentences, a CPS appeared after the comma, just as found with spoken sentences. This effect was also found under poor comma rule knowledge (i.e., pauses in unexpected locations), therefore indicating that more than implicit phonological representations, there is an online processing of these sentence boundaries. This result shows that CPS may be a neural process which structures the whole sentence information (Frazier, Carlson, & Clifton, 2006, Steinhauer & Friederici, 2001). However, not all silent reading tasks elicit a CPS. As Kerkhofs, Vonk, Schriefers and Chwilla (2008) have shown, the CPS seems to be dependent on the engagement of attention towards the comma. It is only when subjects attentively segment the sentence that a CPS is elicited. Due to this reason, the experimental task is critical to elicit a CPS; because Steinhauer used unexpected commas, subjects had to pay additional attention to them, while in Kerkhofs the phrasing was facilitated and guided by syntax. If attention is an important aspect to explain the CPS, then it could be hypothesized that by controlling the amount of attention needed to segment a sentence, one could influence CPS. With this in mind, Li and Yang (2010) examined how prosodic hierarchy influences boundary processing using Chinese poems, a well-structured and rhythmic discourse. These poems have several characteristics: couplet boundaries, which are similar to a paragraph break, intonational phrase boundaries, separating two sentences, and phonological phrase boundaries, or the breaks inside a sentence. Although CPS appeared in the three conditions, its latency differed. Couplet boundary had CPS around 710-720 ms, intonational phrase boundaries around 500-520 ms, and phonological phrase boundaries around 450-470 ms. This difference was interpreted as an interference of the prosodic hierarchal level, in which the lower-level prosodic boundaries had an advantage in processing time. In fact, when processing a complex discourse we must first decode the meaning in an intonational phrase, then integrate that information into the complete sentence, and finally integrate the whole sentence in a series of sentences (Li & Yang, 2010). There is an increase in the information to be processed and it interferes with the timing at which the CPS is elicited. This study reflects how even at a rather basic level of speech perception, the prosodic level, the brain requires additional time to process higher levels of information.

Is CPS present in infants? Pannekamp, Weber and Friederici (2006) reported a positive peak in 8-month-old infants that could be interpreted as a CPS. Thus it appears that the CPS is present even before language acquisition. Interestingly, the latency of this peak is around 2000 ms after the phrase boundary, much later than the typical 300-700 ms latency found in adults. This difference in latency was interpreted as a developmental continuity in underlying neural-basis language processing, in which the neurocognitive process to segment language is still emerging and at a very initial stage (Pannekamp et al., 2006). With more experience and knowledge of language, the CPS becomes faster as a result of implicit knowledge of phonological representations of a given language. However, a recent study by Mannel and Friederici (2011) compared CPS in children with 2, 3 and 6 years of age. The CPS appear in 3 - and 6-year-olds, but not in 2-year-olds. These differences seem to correlate with the age where syntactic rules are learned, around 3 years, allowing children to segment intonational phrases more easily (not only guided by lexical content). These findings indicate that the CPS develops early and is related to gaining implicit knowledge of syntactic rules. Whether the differences between these two studies relate to experimental differences, or different time-windows of analysis, remains to be clarified. A very different, but somehow complementary study was conducted by Steinhauer, Abada, Pauker, Itzhak, and Baum (2010) with elderly participanst, 65 to 80 years of age. Older subjects exhibited a CPS that occured between 0 to 150 ms at the end of the intonational phrase. These results indicate that more experience in processing speech anticipates CPS. However, more research is needed because the latency of the CPS varies according to the experimental paradigm: CPS appears immediately after pause onset, in the 0-300 ms range (Steinhauer & Friederici, 2001; Steinhauer, 2003), and only on the 300-700 ms time window (Li & Yang, 2010, Pannekamp et al., 2005, 2006) for subjects within the same age group. This shows that different stimuli or experimental methods may affect the latency, even when controlling the subject age.

Similarly to language, musical phrases also carry information to allow the listener to parse each melodic phrase. Replicating the experimental methods used in speech CPS, Knosche et al. (2005) found a component with the same characteristics of the speech CPS, but delayed in time. This component, termed CPSm, appears time-referenced to the onset of a second phrase, at 450-600 ms, unlike the speech homologue which is referenced to the end of an intonational phrase, between 300-700 ms. Due to these differences the speech and the music CPS are not strictly the same component, even though they may rely on shared mechanisms to detect phrase boundaries. Neuhaus, Knosche and Friederici (2006) found that musical expertise influences the CPSm, as musicians seem to process the boundaries in music in a similar way to language, yet non-musicians reveal no CPSm (Neuhaus et al., 2006). This difference seems to be related to the experimental paradigm, as when subjects were asked to listen to the music in a holistic manner, CPSm was also found in non-musicians (Nan, Knosche & Friederici, 2009). With the MEG technique, musicians and non-musicians showed a CPS, with musicians having larger amplitudes than non-musicians (Neuhaus et al., 2006). Furthermore, the study by Nan et al. (2009) reported differences in CPSm latency according to cultural background (100 ms advantage for same culture music) and musical expertise (50 ms advantage). An interesting fact about the comparison between the speech CPS and music CPS is that both vary in latency depending on different types of variables. Differences in latency could be due to an highly competent neural mechanism to segment information, which depends on the type and characteristics of both the listener and the stimuli. For example, the CPS in 8-month-old infants shows a significant delay in its onset. When hearing music, there is also a delay on the onset of CPS. We might speculate that the harder it is segment prosodic boundaries, the longer the latency irrespective of domain.

In sum, the CPS appears to reflect a neural a neural network responsible for chunking prosodic information. Its latency depends on prosodic hierchical level (Li & Yang, 2010), type of stimulus (Pannekamp et al., 2005) and also subject attributes such as age and language proficiency (Pannekamp et al., 2006; Steinhauer et al., 2010). Its topography varies according to type of stimulus, with less lexical information associated with anterior topography (Pannekamp et al., 2005). The CPS is elicited in auditory (Steinhauer & Friederici, 2001) as well as visual modality (Steinhauer, 2003). Also, differences in amplitude and latency have been found between musicians and non-musicians in CPSm (Nan et al., 2009, Neuhaus et al., 2006). To our knowledge no study has examined these differences in the speech CPS. That is the goal of the present study. Studies on the CPS so far have not investigated how musical expertise may affect boundary processing in speech. Based on evidence of effects of musical expertise on other ERP components related to speech processing (Marques et al., 2007, Moreno et al. 2009), and on the music CPS (Nan et al., 2009, Neuhaus et al., 2006), we set out to investigate how the speech CPS is affected by musical expertise. Our prediction is that the speech CPS would be modulated by musical expertise, namely its latency and topography. To do so, we will compare musicians and non-musicians in a probe detection paradigm using two types of stimuli: spoken and hummed sentences. In the former, boundary processing may be guided by prosodic, lexical and syntactic cues, whereas in the second boundary processing has to rely on prosodic cues alone.

2. Method

Participants

Thirty-two participants were recruited to perform the experiment, which lasted 45 minutes (experimental blocks), plus 30 minutes for the EEG setup. All participants received a small token of appreciation, either formation credits or a small monetary contribution. Half of the subjects (sixteen) were musicians (9 women, mean age: 20.2 years old, $SD = 3$) and half non-musicians (10 women, mean age: 20 years old, $SD = 3.5$). All musicians had at least 7 years of formal training ($M = 10.8$ years, $SD = 2.68$) and a minimum of two hours of weekly practice ($M = 6.25$, $SD = 6$). All musicians started musical training before 10 years of age ($M = 6.81$, $SD = 2.25$). Non-musicians had no musical training, except those of exogenous classes in elementary school. None of the subjects declared having sight problems, hearing difficulties, neurological or neuropsychiatric disorders or being currently under any medication. All subjects were Portuguese natives, and all were right-handed accordingly to the Edinburgh Handedness Test (Oldfield, 1971).

All subjects were asked to read an explanation of the experiment (see appendix 1), and sign an informed consent before the data retrieval (see appendix 2). Subjects were given a standardized set of instructions (see appendix 3), and shown the effects of physical movements on the EEG data. They were further instructed to avoid moving, to stay as relaxed as possible and to avoid blinking. Finally, they were told they could abandon the experiment at any given time, if they did not feel comfortable (none did). This study was performed under Helsinki declaration rules for good research practice. In the end of the experiment, subjects were asked to leave their contact for the debriefing and further input on the results from their data.

Stimuli

For the experiment we selected 48 sentences. These sentences would then be adapted into two groups – one with one phrase boundary (condition A) and another with two (condition B), using roughly the same words (word list can be seen in appendix 5). The selection of the words and sentences underwent a rigorous control of psycholinguistic variables with resource to Porlex (Gomes & Castro, 2003). All sentences were declarative and were composed of different types of syntactic constituents. These include coordinating and subordinating phrases. Connectives such as “and” and “but”, as well as the conjunctions “though” and “if” were used, as exemplified below.

Example 1: Coordinated phrases using enumeration (“e”, and)

A: O João comprou carne [IPh1], o Jorge e a Luísa trouxeram saladas e bebidas [IPh2].

John bough meat, Jorge and Luise brought salads and drinks

B: O João comprou carne [IPh1], o Jorge trouxe saladas [IPh2], e a Luísa trouxe bebidas [IPh3]. John bough meat, George brought salads, and Louise brought drinks

Example 2: Coordinated phrases with “but” (mas)

A: O pai adorou a peça [IPh1], mas a mãe e os filhos não gostaram mesmo nada [IPh2].

The father loved the play, but the mother and sons did not like it at all.

B: O pai adorou a peça [IPh1], mas a mãe detestou [IPh2], e os filhos não gostaram nada [IPh3]. The father loved the play, but the mother hated it, and the sons did not like it at all.

Example 3: Subordinating phrase with “although” (embora), main clause second

A: Embora esteja frio [IPh1], já se sente um calorzinho do sol e um ar leve de Verão [IPh2].

Although it is cold, it feels a bit of heat and a fresh summer air.

B: Embora esteja frio [IPh1], o sol já está brilhante [IPh2], e sente-se um ar leve de Verão [IPh3]. Although it is cold, the sun is already bright, and one feels a fresh summer air.

Example 4: Subordinating phrase with “only if” (desde que), main clause first

A: Concordei com tudo [IPh1], desde que pudesse ver e também experimentar por um dia [IPh2].

I agreed with everything, only if I could see and also try it for one day

B: Concordei com tudo [IPh1], desde que pudesse ver [IPh2], e depois experimentar por um dia [IPh3] I agreed with everything, only i I could see, and try it for one day.

Sentences were recorded in the sound-attenuated booth of the Speech Laboratory at the University of Porto (psychology department). All stimuli were recorded by a Portuguese native female speaker, with musical training including singing lessons and overall experience in recording emotional prosody stimuli (e.g. Castro & Lima, 2010). Each sentence was produced and recorded in normal and hummed speech. In order to achieve an accurate rendering of the hummed sentences, the speaker was instructed to produce the sentences keeping the words in mind and "translating" them into fluent "mmmm" sentences (that is, to literally hum the sentences as meaningful utterances).

Sentences were digitized with a sampling rate of 48 kHz and 24-bit resolution using Pro Tools LE, version 5.1.1, and a high quality microphone attached to a Macintosh G4 computer. The acoustical files were saved in AIFF format and posteriorly converted to WAVE, to suit the stimulation software. Using Soundforge, each soundfile was inspected at a 32:1 zoom, and a silence of 20 ms was inserted before the first word onset. Amplitude was normalized at 0 dB peak.

For further acoustical analysis, we selected critical periods for each sentence, namely intonational phrase 1 (IPh1), pause 1 and intonational phrase 2 (IPh2) for condition A, and IPh1, pause1, IPh2, pause2 and intonational phrase 3 (IPh3) for condition B. Each part was measured using Praat 4.6.36, and mean, min and maximum duration of these parts can be seen in table 1 and 2, respectively. Importantly, we found no statistical differences between the first intonational phrase, between pause duration in the first and second boundary, and pause duration between condition A and B. This shows that stimuli was equivalent in the first part of the sentence (as intended), and different in the second (phrased versus unphrased, also as intended).

Table 1. Mean duration in milliseconds of the critical parts of sentences with 1 (A) or 2 (B) boundaries, in spoken and hummed versions. SD is also shown.

	IPh 1	Pause 1	IPh 2	Pause 2	IPh 3	Total
A spoken	1119 ± 119	354 ± 101	2692 ± 220	na	na	4165 ± 278
B spoken	1163 ± 136	379 ± 114	1211 ± 137	376 ± 116	1457 ± 152	4586 ± 264
A hummed	1142 ± 81	471 ± 105	3129 ± 294	na	na	4742 ± 366
B hummed	1219 ± 110	621 ± 110	1277 ± 146	594 ± 115	1648 ± 251	5359 ± 344

Table 2. Minimum and maximum duration of the critical parts of sentences with 1 (A) or 2 (B) boundaries, in spoken and hummed versions

	IPh 1	Pause 1	IPh 2	Pause 2	IPh 3	Total
A spoken	881 - 1474	120 – 576	2200 – 3328	na	na	3465 - 4888
B spoken	874 - 1552	160 – 612	960 – 1504	160 - 672	1108 - 1808	3647 - 5167
A hummed	976 - 1370	232 – 712	2339 – 3647	na	na	3846 - 5533
B hummed	980 - 1475	288 – 792	1008 – 1600	360 - 856	1088 - 2379	4754 - 6198

The difference of the fundamental frequency in sentences was also analyzed. We computed mean F_0 at corresponding points in time for all sentences in the same condition. Intervals of 10 ms were chosen from the onset of the sentences onwards, for a period equal to the shortest soundfile in each condition. This was 3400 ms for A spoken, 3600 ms for B spoken, 3800 ms for A hummed, 4700 ms for B hummed. As seen, hummed and spoken sentences have fairly similar pitch trackings at the fundamental frequency, across all sound files (Figure 1).

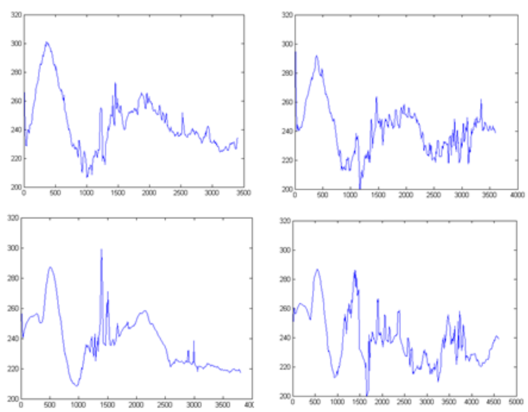

Figure 1. Mean F_0 for 10 ms intervals from sentence onset, for spoken A (upper left) and B sentences (upper right) and for hummed A (lower left) and B sentences (lower right).

We also analysed pitch movement at critical segments. A descending F_0 pattern was expected for utterance final prosodic phrases (Wang & Hirschbergt, 1992). We plotted the F_0 curves of critical segments, which are the end of IPhs. A time window of 500 ms was chosen, that is, the preceding 500 ms of IPh offset. (Figure 2). As seen, there is a descending pattern before the end of the intonational phrase both in hummed and spoken conditions.

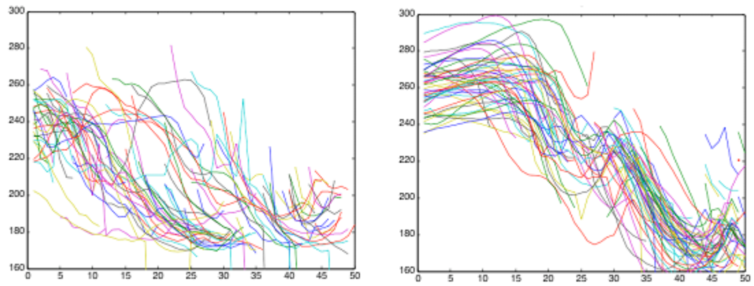

Figure 2. F_0 plots of all 46 files for 500 ms pre-offset of final phrases in sentences with two boundaries (IPh2), spoken (left panel) and hummed (right panel).

Procedure

Subjects sat in a comfortable chair, roughly 1,5 meters from a 17 inch LCD computer screen, and were given documents regarding the experiment, namely the informed consent form, the experiment information and a custom-made questionnaire for data collection (please see appendix 4). The applicant would use the Edinburgh handedness questionnaire to test subject's laterality. A brief overview of the EEG montage procedure was given orally. Subjects were instructed to focus their sight on a cross at the center of the screen, and to answer whether a probe word was present in the heard sentence. They did so, by answering "yes", or "no", in an input controller, with left and right buttons designed for the experiment. Stimuli were presented through high-quality headphones, while probe words were presented visually on the monitor.

Stimuli were divided into 3 blocks for both hummed and spoken sentences, in a total of 6 experimental blocks. Rest periods would occur between each block. Before the first block of spoken or hummed sentences, a training block was performed to aid in task understanding. Each spoken block had 30 sentences that lasted for roughly 5 minutes, while hummed blocks had 30 sentences and 15 fillers that lasted an average of 8 minutes. Block order was pseudo-randomized across subjects. Probe words would be correct on a 50% chance level. For hummed conditions, due to the lack of lexical information, 45 additional stimuli were created, where a word was artificially spliced into the acoustical stream. These stimuli were accounted as fillers made to preserve the experimental paradigm, and were unaccounted in further analysis. An example of a trial can be seen in Figure 3.

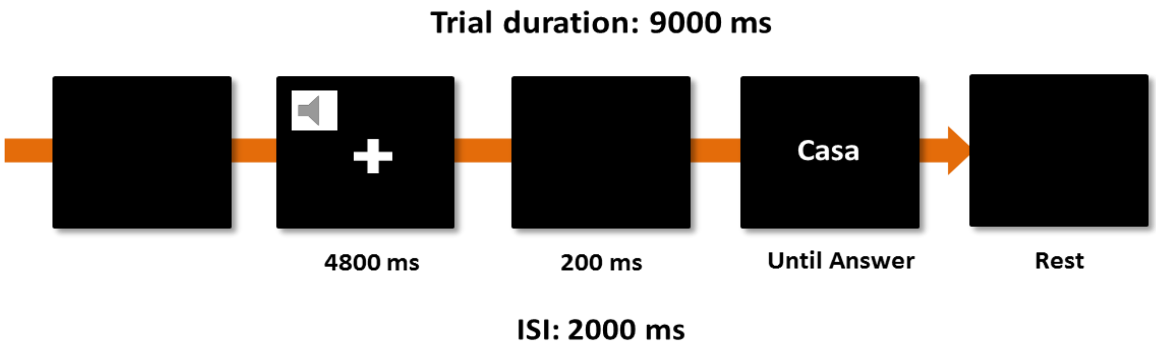

Figure 3. Example of an experimental trial. After a blank screen, a sentence was presented auditorily while a cross was presented at the center of the screen. After 200 ms, a probe word was presented until the onset of response. Subjects decide whether the word was present in the sentence. During the inter-stimuli interval of 2000 ms, the screen would be blank.

The experimental design would then be completed by adding triggers into critical parts of the sentence. The trigger and stimulation software used was Presentation. Based on previous experiments by Pannekamp et al. (2005), we selected 5 points of interest, namely, the start (T1) and end (T2) of IPh1, the start (T3) and end (T4) of IPh2, and finally, the start (T5) of the IPh3 on condition B, and the exact latency for condition A where there is no IPh3 (therefore, continuous speech). The first three triggers would be comparable across conditions, while the final trigger would compare a boundary versus continuous speech. An example can be seen in Figure 4.

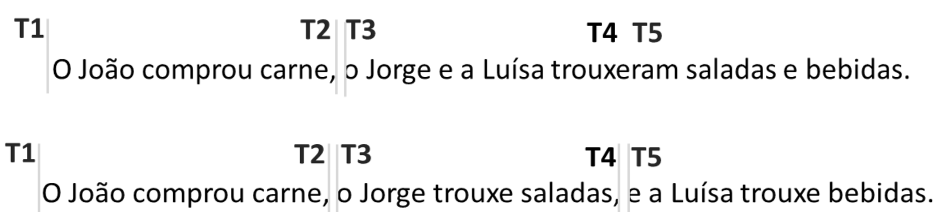

Figure 4. Visual representation of the trigger points for time synchronization, condition A with one boundary (top) and condition B with two boundaries (bottom).

ERP procedure

EEG data was recorded in a sound-attenuated booth, using ASA 4.6 software, a 32 channel Ag/AgCI cap, and a waveguard REFA-32 amplifier. A sampling rate of 512 Hz was selected. We replicated Pannekamp et al. (2005) methods for a good comparison between experiments. Therefore, data was recorded using left mastoid as reference, and was offline re-referred to both mastoids. During the recording, some subjects had a band-stop filter at 50 Hz, due to electrical interference. No additional filters were applied online. An EOG channel was placed beneath the subject's right eye, to better detect blinking artifacts. Furthermore, lateral electrodes were placed on the left and right external canthi, to account for eye movements.

EEG data was processed using ASA 4.8 software. A band-pass filter of 0.01 to 30 Hz (12 dB/Oct) was used. Artifact detection was done by maximum values exclusion, with a criteria of 80 microvolts exceeding baseline activity. Furthermore, all trials were visually inspected and trials with eye movements, blinks, alpha waves, muscle activity and general noise were removed from further data processing (mean = 30%). Due to excessive artifacts, two subjects were removed from musicians, and one from non-musicians.

In trigger 4, baseline correction was performed using 200 ms pre-stimuli, aligned with the trigger point. Due to the presence of exogenous components from the next sentence onset, we also performed analysis on trigger 5 to measure P50, N100 and P200 without latency jitter. In trigger 5, we selected a baseline on -800 to -600 ms for hummed sentences. Due to the short pause duration on spoken sentences we did not analyze the equivalent time-window for this condition, as it was expected to be contaminated with previous ERP components.

The data was averaged across epochs using the whole trial duration, and then across all subjects for the grand average. For CPS trigger 4 data analysis, consecutive epochs of 100 ms were computed from 0 to 1200 ms after the end of the intonational phrase. For latency analysis, we used a single time-window from 0 to 1200 ms. In trigger 5, we computed -800 to 600 ms time-windows for hummed sentences and -600 to 600 ms for spoken sentences. For exogenous component measures we selected critical time-points based on visual inspection. For CPS preceding the trigger 5 in hummed condition, we selected a 300 ms time-window from -300 to 0 accordingly to pause duration. The computed measures were max latency for latency analyses, and average amplitude for the remaining events of interest.

For source localization we used sLORETA (standardized low-resolution brain electromagnetic tomography; Pasqual-Marqui, 2002) with 20 mm grid spacing, in conjunction with a standard MNI headmodel, electrodes and a Tailaraich adapted MRI file. For anatomical localization, we used the brain atlas from Tailaraich Client 2.4.2 software. All deep (subcortical) sources were disregarded from further analysis as the limitations on source localization using EEG make such results unlikely.

Statistical analysis

All statistical analysis were performed using Statistica 10.0.1. Analysis of variance (ANOVA) was used for all statistical tests, and all values were adjusted with Greenhouse-Geisser epsilon correction for non-sphericity. Repeated measures ANOVAs were computed separately for midline electrodes and regions of interest. For midline, factors topography (FZ, CZ and PZ), group (musicians and non-musicians) and phrasing (phrased and unphrased) were computed. For region of interest analysis, an additional hemisphere (left and right) factor was added, crossed with topography (2 x 3). Each ROI included three electrodes: frontal left (F7, F3, Fc5), frontal right (F8, F4, Fc6), central left (C3, Cp1, Cp5), central right (C4, Cp2, Cp6), parietal left (P7, P3, O1), and parietal right (P8, P4, O2). Whenever topography factor showed significant interactions, further ANOVAs were conducted separately for each region. Tukey tests were used for all post-hoc analysis. Topographical maps were computed using ASA 4.8 software. All subjects reached more than 90% correct responses. Given the irrelevance of the behavioral task for CPS, behavioral data were not further analyzed.

3. Results

Analysis by end of intonational phrase (trigger 4)

Spoken sentences at midline

Musicians mean amplitude differed from non-musicians in both early and latter time-windows (100-200, 300-400 and 800-900 ms; see Figure 5). CPS was found in the 200 to 900 ms time-window, as observed by the differences between phrased and unphrased conditions. Three-way repeated measures found significant interactions of *Phrasing* (P) x *Group* (G) x *Topography* (T) in the 100-200 ms time-window [$F(2, 54)$, $p < .01$, $\eta^2 = .20$]. Additional analysis revealed a significant interaction between F x G [$F(1, 26) = 7.72$, $p < .01$, $\eta^2 = .22$] on frontal topography. Tukey post-hoc showed that musicians ($M = 1.74$ µV, $SD = 0.69$) differed from non-musicians ($M = -0.67$ µV, $SD = 0.67$; $p < .05$) when a phrase boundary was presented. Furthermore, non-musicians phrased condition was significantly different from the unphrased counterpart ($p < .05$). In 300-400 ms, a P x G x T interaction was close to significance [$F(2, 56) = 2.96$, $p = .089$, n.s.]. In time-window 300-400 ms, an interaction of P x G was found [$F(1, 26) = 6.24$, $p < .05$, $\eta^2 = .19$]. Again, post-hoc analysis showed that musicians ($M = 3.88$ µV, $SD = 0.57$) had larger amplitudes than non-musicians ($M = 1.55$ µV, $SD = 0.54$; $p < .01$). From time-windows 400 to 800, no significant differences were found between musicians and non-musicians, although main effects of phrasing were present in all conditions (see table 1). Close to significant main interactions of group were found in time-windows from 500 to 800 (*ps* < .08, ns.). The time-window 800-900 ms showed a significant interaction of P x G x T [$F(2, 56) = 6.39$, $p < .01$, $\eta^2 = .19$]. Two-way ANOVAs across topography showed an interaction between phrasing and group in both frontal [$F(1, 27) = 18.71$, $p < .01$, $\eta^2 = .40$] and central regions [$F(1, 27) = 11.43$, $p < .01$, $\eta^2 = .29$]. Post-hoc analysis in frontal topography shows that non-musicians had differences in phrasing, while musicians did not. In central areas, musicians showed larger amplitudes in phrased conditions ($M = 1.28$ µV, $SD = 0.23$) when compared to unphrased conditions ($M = -0.03$ µV, $SD = 0.32$; $p < .05$). Finally, in time-window 900-1000 ms, a P x T interaction was found [$F(1, 26) = 8.56$, $p < .01$, $\eta^2 = .24$]. Tukey post-hoc showed that CPS would still appear at this latency, however, only in frontal regions, as confirmed by the phrased condition ($M = 1.06$ µV, $SD = 0.24$) being larger than the unphrased counterpart ($M = 0.18$ µV, $SD = 0.26$; $p < .01$).

Table 3. Results of the ANOVAs for midline and lateral electrodes in successive latency bands on spoken sentences

Latency (ms)	Electrode	Interaction	Frontal	Central	Parietal
0-100	Midline	-	-	-	-
	Lateral	-	-	-	-
100-200	Midline	P x T x G: $F_{2,54}$ = 6.95**	F x G: $F_{1,27}$ = 7.72**	-	-
	Lateral	P x T x G: $F_{2,54}$ = 5.64*	F x G: $F_{1,27}$ = 6.64*	-	-
200-300	Midline	-	-	-	-
	Lateral	-	-	-	-
300-400	Midline	F x G: $F_{1,27}$ = 6.24*	-	-	-
	Lateral	F: $F_{1,27}$ = 11.48**	-	-	-
400-500	Midline	F: $F_{1,27}$ = 17.42***	-	-	-
	Lateral	F x T x H: $F_{2,54}$ = 3.86*	-	F: $F_{1,28}$ = 11.06**	F x H: $F_{1,28}$ = 21.87***
500-600	Midline	F: $F_{1,27}$ = 24.32 ***	-	-	-
	Lateral	F x T x H: $F_{2,54}$ = 8.03**	F: $F_{1,27}$ = 4.81*	F: $F_{1,27}$ = 21.84***	F x H: $F_{1,28}$ = 14.88***
600-700	Midline	F: $F_{1,26}$ = 14.61***	-	-	-
	Lateral	F: $F_{1,26}$ = 8.51**	-	-	-
700-800	Midline	F: $F_{1,26}$ = 8.17**	-	-	-
	Lateral	F x H: $F_{1,28}$ = 7.69**	-	-	-
800-900	Midline	P x T x G: $F_{2,54}$ = 6.39*	F x G: $F_{1,27}$ = 18.71***	F x G: $F_{1,27}$ = 11.43**	-
	Lateral	F x T: $F_{2,54}$ = 7.27**	-	-	-
900-1000	Midline	F x T: $F_{2,54}$ = 8.36	-	-	-
	Lateral	F x T: $F_{2,54}$ = 9.26**	-	-	-
1000-1100	Midline	-	-	-	-
	Lateral	-	-	-	-
1100-1200	Midline	-	-	-	-
	Lateral	-	-	-	-

Note: differences between conditions were significant at * = .05; ** = .01; *** = .001.

Spoken sentences by Regions of Interest

Similarly to midline analysis, musicians differed from non-musicians in early and latter time-windows. In time-window 100-200 ms, an interaction between F x T x G was found [$F(2, 56) = 5.64$, $p < .05$, $\eta^2 = .17$]. Further analysis revealed that musicians (M = 0.62 µV, SD = 0.48) had larger amplitudes than non-musicians (M = -1.18 µV, SD = 0.46; $p < .05$) in both hemispheres, while non-musicians also were different in phrased (M = -1-18 µV, SD = 0.46) and unphrased (M = 1.11 µV, SD = 0.37) conditions ($p < .01$). These results were only found in frontal regions. From time-window 300 to 1000 ms, all conditions elicited a CPS (see table 1). Furthermore, an effect of laterality was found in this time period. Specifically, in time window 700-800, CPS was only significant in the left hemisphere confirmed by the interaction of P x H [$F(1, 27) = 7.69$, $p < .01$, $\eta^2 = .22$] where the left hemisphere (M = 1.22 µV, SD = 0.25) showed larger amplitude than right hemisphere (M = 0.37 µV, SD = 0.15; $p < .01$). Furthermore, in 800-1000 ms time windows, CPS was only found in frontal regions (see Figure 6 for a topographical overview).

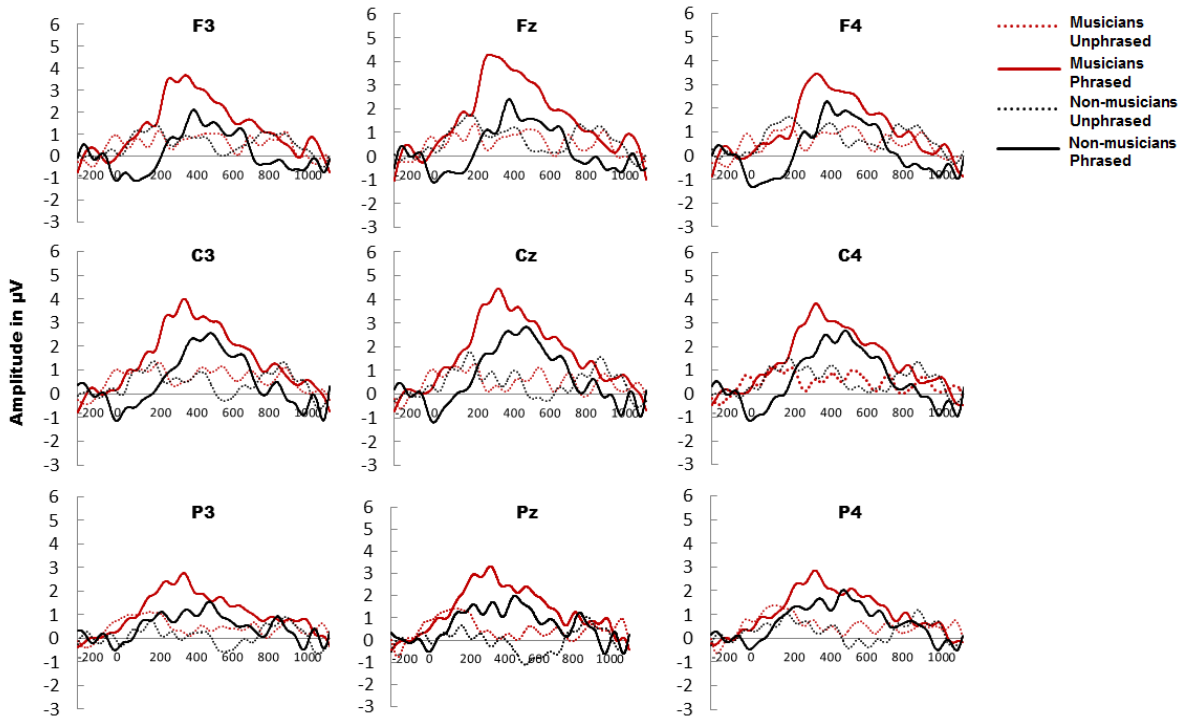

Figure 5. ERPs (averaged mean over frontal, central and parietal electrode sites) for Spoken condition for the time window -200 to 1200 ms. Musicians in red, non-musicians in black. Dotted lines represent unphrased and solid lines phrased. Positive voltage is plotted up.

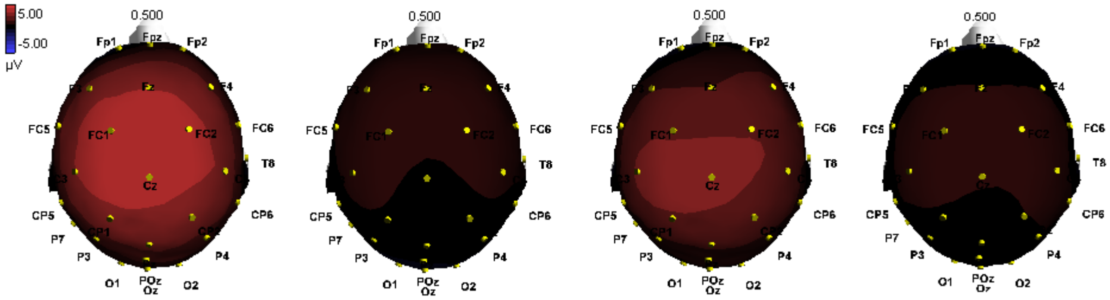

Figure 6. Topographic maps at 500 ms in spoken sentences for musicians and non-musicians. From left to right: musicians phrased, musicians unphrased, non-musicians phrased, non-musicians unphrased. Voltage in microvolts and color-coded, red for positive and blue for negative.

Hummed sentences at midline

In hummed midline analysis, time-window 200-300 ms showed a main effect of group [$F(1, 27) = 4.24$, $p < .05$, $\eta^2 = .13$], in which musicians have overall larger amplitudes than non-musicians. From 200 to 1000 ms, CPS was present in both musicians and non-musicians (see table 4). However, in 200 to 300 ms, CPS was only found in frontal areas, as seen by the P x T interaction [$F(2, 54) = 13.38$, $p < .01$, $\eta^2 = .33$]. All other time periods showed larger amplitudes in both Frontal and Central topographies (all *ps* < .05; see figure 7), but not in parietal areas. Again, a main interaction of Group was almost significant at 600-700 ms [$F(1, 27) = 4.07$, p = .053].

Hummed sentences by Regions of Interest

Analysis by regions of interest showed that musicians had overall larger CPS than non-musicians. Also, this effect seems to be lateralized to the left hemisphere. In 100-200 time-window, P x T, H x G and F x T interactions were significant. However, after multiple comparisons using Tukey, the only consistent effect was the H x G interaction [$F(1, 27) = 4.86$, $p < .05$, $\eta^2 = .15$], with musicians ($M = -0.31$ μV, $SD = 0.27$) having lower values than non-musicians ($M = -0.01$ μV, $SD = 0.22$; $p < .05$). Similarly, in 200-300 ms, we found an interaction of F x G x T [$F(2, 56) = 2.49$, $p < .05$, $\eta^2 = .19$] Further analysis by topography showed a main effect of Phrasing [$F(1, 27) = 8.82$, $p < .01$, $\eta^2 = .25$] and a main effect of Group [$F(1, 27) = 4.39$, $p < .05$, $\eta^2 = .15$] in frontal electrodes, independently of hemisphere (see Figure 8 for a topographical overview).

Table 4. Results of the ANOVAs for midline and lateral electrodes in successive latency bands on hummed sentences

Latency (ms)	Electrode	Interaction	Frontal	Central	Parietal
0-100	Midline	-	-	-	-
	Lateral	-	-	-	-
100-200	Midline	-	-	-	-
	Lateral	H x G: $F_{1,27}$ = 4.86*	-	-	-
200-300	Midline	F x T: $F_{1,27}$ = 13.38*** G: $F_{1,27}$ = 4.29*	-	-	-
	Lateral	-	-	-	-
300-400	Midline	F x T: $F_{2,54}$ = 11.98**	-	-	-
	Lateral	F x T: $F_{2,54}$ = 33.24***	-	-	-
400-500	Midline	F x T: $F_{2,54}$ = 35.1***	-	-	-
	Lateral	F x T: $F_{2,54}$ = 26.71***	-	-	-
500-600	Midline	F x T: $F_{2,54}$ = 42.23***	-	-	-
	Lateral	F x T x H x G: $F_{2,54}$ = 27.08**	F x H x G: $F_{2,54}$ = 7.21*	F: $F_{2,54}$ = 33.29***	F: $F_{2,54}$ = 5.78*
600-700	Midline	F x T: $F_{2,54}$ = 27.08**	-	-	-
	Lateral	F x T x H x G: $F_{2,54}$ = 6.1**	F x H x G: $F_{2,54}$ = 6.37**	F x H: $F_{2,54}$ = 5.14**	F: $F_{2,54}$ = 5.03*
700-800	Midline	F x T: $F_{2,54}$ = 14.45***	-	-	-
	Lateral	F x H x G: $F_{2,54}$ = 6.05	-	-	-
800-900	Midline	F x T: $F_{2,54}$ = 11.05**	-	-	-
	Lateral	F x T: $F_{2,54}$ = 11.65***	-	-	-
900-1000	Midline	F x T: $F_{2,54}$ = 10.38***	-	-	-
	Lateral	F x T x H x G: $F_{2,54}$ = 6.59**	-	H x G: $F_{2,54}$ = 6.03**	F x H x G: $F_{2,54}$ = 4.26**
1000-1100	Midline	F: $F_{2,54}$ = 7.79**	-	-	-
	Lateral	F x H x T: $F_{2,54}$ = 3.7*	F: $F_{2,54}$ = 8.4**	F: $F_{2,54}$ = 7.11*	F x H: $F_{2,54}$ = 4.26**
1100-1200	Midline	F x T: $F_{2,54}$ = 5.45*	-	-	-
	Lateral	F x T: $F_{2,54}$ = 10.18**	-	-	-

Note: differences between conditions were significant at * = .05; ** = .01; *** = .001.

In central electrodes, an P x H interaction was found [$F(1, 27) = 4.93$, $p < .05$, $\eta^2 = .14$], with Tukey showing that phrased conditions (M = 0.29 µV, SD = 0.29) had larger amplitudes than unphrased conditions (M = -0.32 µV, SD = 0.27; $p < .05$), but only in the left hemisphere. Time-windows 300 to 500 ms showed a CPS but only in frontal and central topographies. Additionally, an interaction of P x H x G x T was found from 500 to 600 ms [$F(2, 54) = 3.46$, $p < .05$, $\eta^2 = .11$], and from 600 to 700 [$F(2, 54) = 6.1$, $p < .01$, $\eta^2 = .11$]. Further two-way ANOVAs revealed similar effects for both time-windows: in frontal regions, both hemispheres showed main effects of phrasing, but only left hemispheres had group effects. In central and parietal regions, both hemispheres showed a CPS and no group effects (see table 1). Similarly, in 700 to 800 ms, an interaction of P x H x G was found [$F(2, 54) = 6.05$, $p < .05$, $\eta^2 = .18$]. Further analysis showed group effects for both hemispheres, where musicians had larger values than non-musicians independently of phrasing. The remaining time-windows showed Phrasing effects, but only in frontal and central regions, while parietal regions would also differ significantly, but only from 900 to 1100 ms, and in right hemispheres.

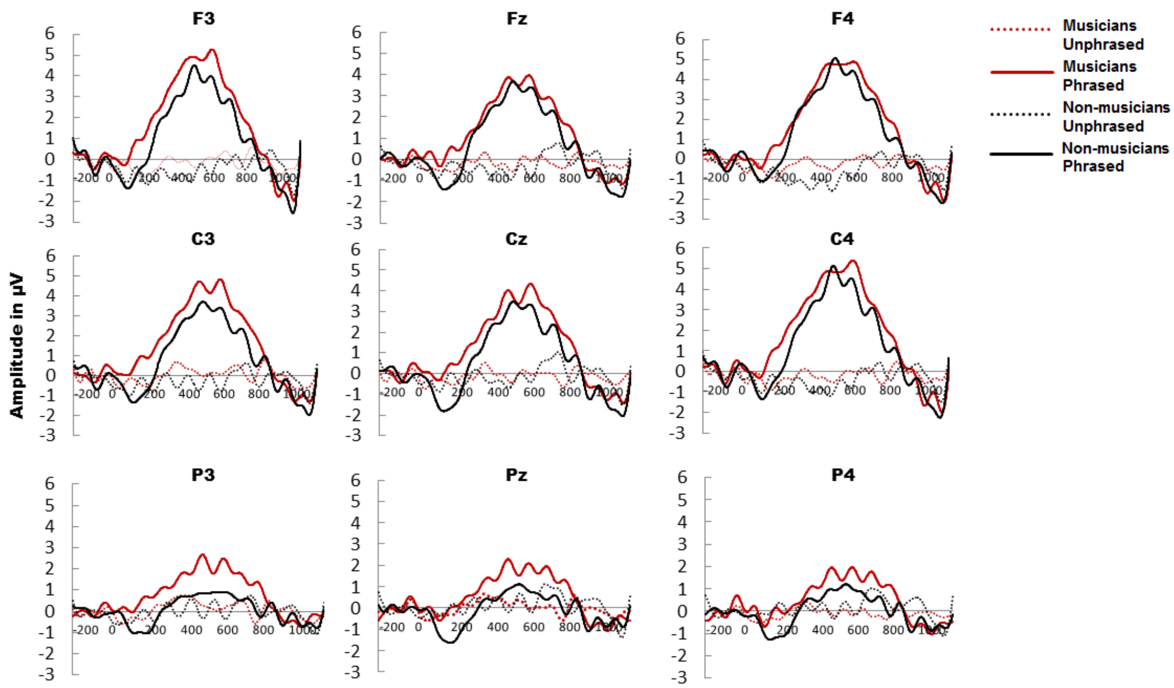

Figure 7. ERPs (averaged mean over frontal, central and parietal electrode sites) for hummed condition in the time window -200 to 1200 ms. Musicians in red, non-musicians in black. Dotted lines represent unphrased and solid lines phrased. Positive voltage is plotted up.

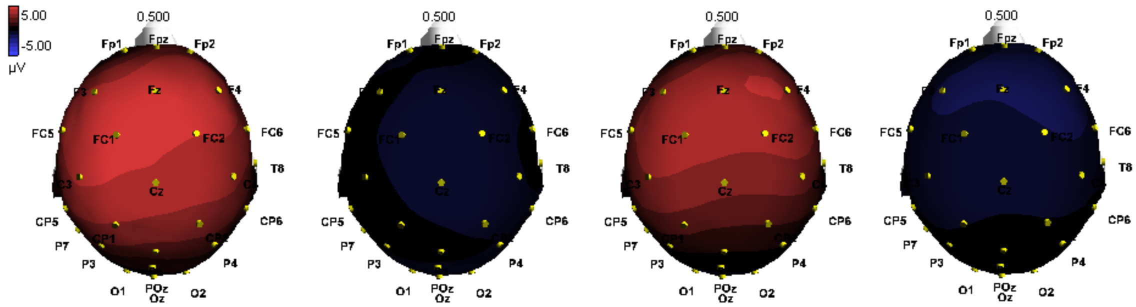

Figure 8. Topographic maps at 500 ms in hummed sentences for musicians and non-musicians. From left to right: musicians phrased, musicians unphrased, non-musicians phrased, non-musicians unphrased. Voltage in microvolts and color-coded, red for positive and blue for negative.

Latency analysis

For latency analysis, max amplitude values were computed trough ASA 4.8 software. This algorithm retrieves the time on which the peak appears at maximum value. Time-windows 0-1200 ms after the intonational phrase offset were selected. Presented means are in respect to midline electrodes.

Spoken sentences

In spoken sentences, musicians (M = 430.7 ms, SD = 50.2) and non-musicians (M = 500.4 ms, SD = 48.5) showed no significant differences. In midline electrodes, a two-way mixed ANOVA revealed no significant interactions or main effects. Most notably, Group effect was not significant [$F(1, 27) = 0.99$, ns]. In regions of interest, three-way repeated measures ANOVA revealed no further significant interactions for hemispheres or topography (all $ps > .05$).

Hummed sentences

Similar to the previous results, we found no differences in latency between musicians (M = 535.7 ms, SD = 35) and non-musicians (M = 564.1 ms, SD = 33.9). In midline analysis, no effects were found. Group main effect was again non-significant [$F(1, 27) = 0.34$, ns]. In regions of interest, the same pattern occurred (all $ps > .05$).

Analysis by intonational phrase onset (trigger 5)

For exogenous components analysis, we computed the mean amplitude for each component with 50 ms in between. Therefore, P50 was measured from 0-100 ms, N100 from 50 to 150 ms and P200 from 150 to 250 ms. Figure 9 shows the average means for all spoken conditions and Figure 10 for hummed conditions.

Exogenous components in spoken sentences

P50 component

For midline analysis, a two-way repeated measures ANOVA revealed no significant effects, including no effects of group (F(1, 28) = 1.19, ns.). In ROI analysis, we found an H x G interaction [$F(1, 28) = 4.87$, $p < .05$, $\eta^2 = .15$), however, Tukey post-hoc analysis shows no significant effects, therefore the effect is discarded.

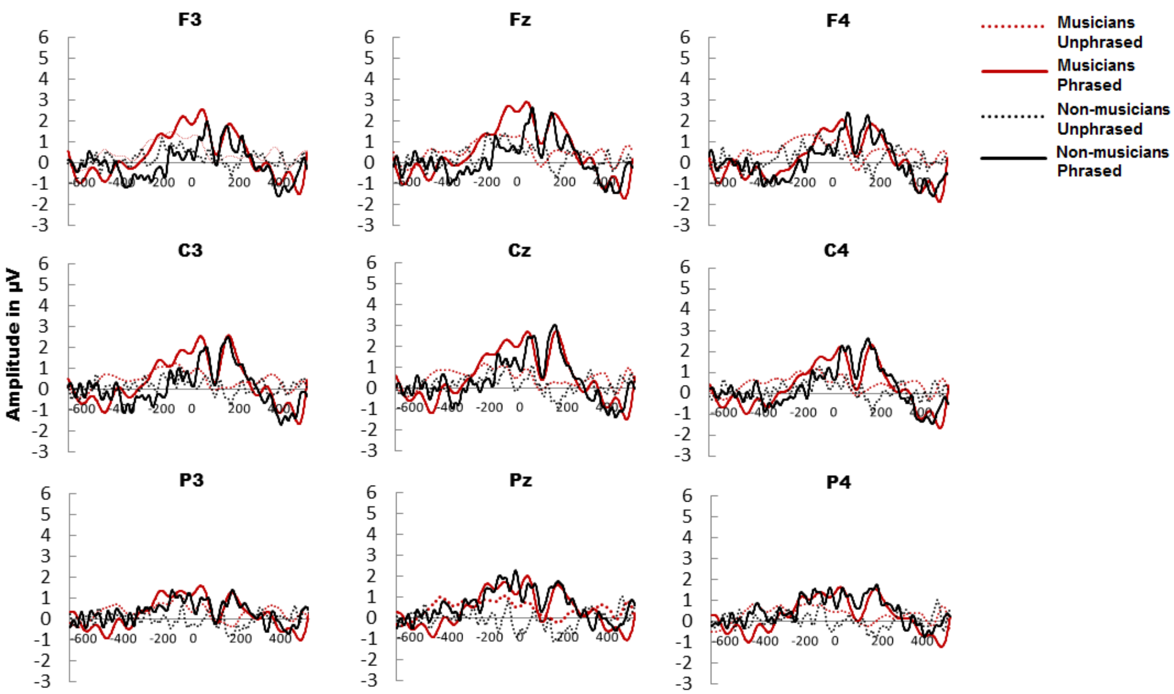

Figure 9. ERPs (averaged mean over frontal, central and parietal electrode sites) for spoken condition for the time window -600 to 600 ms. Musicians in red, non-musicians in black. Dotted lines represent unphrased and solid lines phrased. Positive voltage is plotted up.

N100 component

In midline analysis, no interactions or main effects were found. Group factor showed again no effects [$F(1, 28) = 0.01$, ns.]. In ROI analysis, we found a significant interaction between H x T x G [$F(2, 54) = 3.94$, $p < .05$, $\eta^2 = .13$]. In frontal regions, a significant interaction was found between hemisphere and group [$F(1, 27) = 5.68$, $p < .01$, $\eta^2 = .25$]. Again, after Tukey multiple comparisons, no significant effects were found. In non-musicians, left and right hemispheres were close to significance (p = .075). In central regions, an H x G interaction was also found [$F(1, 27) = 4.45$, $p < .05$, $\eta^2 = .14$]. Post hoc analysis showed no further significant effects. Finally, in parietal regions, a main effect of hemisphere was found [$F(1, 27) = 8.04$, $p < .01$, $\eta^2 = .23$], where the right hemisphere (M = 0.73, SD = 0.32) had larger values than the other half (M = 0.39, SD = 0.33). Musicians and non-musicians had no differences in N100 amplitude [$F(1, 27) = 0.03$, ns.].

P200 component

As for P200 component, we found a main effect of topography in the midline [$F(2, 54) = 4.36$, $p < .05$, $\eta^2 = .14$], with frontal regions showing larger values than central and parietal regions. No main effects of Group were observed [$F(1, 27) = 0.02$, ns.) . In ROI analysis, we found an H x G interaction [$F(1, 27) = 5.3$, $p < .05$, $\eta^2 = .16$). Post-hoc analysis show that non-musicians have larger amplitudes between the right (M = 1.46, SD = 0.28) and left hemisphere (M = 0.93, SD = 0.27; $p < .01$). No effects were found between musicians and non-musicians (all *ps*. > .05)

Exogenous components in hummed sentences

P50 component

For midline analysis, a two-way repeated measures ANOVA revealed a main effect of Topography [$F(2, 54) = 54.49$, $p < .001$, $\eta^2 = .67$), where frontal electrodes showed larger values than parietal electrodes. No effect of group was found [$F(1, 28) = 0.72$, ns.]. In ROI analysis, we found an H x T interaction [$F(2, 54) = 4.89$, $p < .05$, $\eta^2 = .15$], where again frontal electrodes would appear with larger amplitudes than parietal, with a slight shift to the left hemisphere. As with midline analysis, there was no difference between musicians and non-musicians [$F(1, 28) = 0.13$, ns.].

N100 component

In midline analysis, a main effect of topography was found [$F(2, 54) = 55.4$, $p < .001$, $\eta^2 = .67$]. Group factor showed no effects [$F(1, 28) = 0.13$, ns.]. In ROI analysis there were also no differences between musicians and non-musicians [$F(1, 28) = 0.02$, ns.].

P200 component

As for P200 component, we found a main effect of topography in the midline [$F(2, 54) = 21.43$, $p < .001$, $\eta^2 = .44$], with frontal regions showing larger values than central and parietal regions. Again, no effects of Group were found [$F(1, 27) = 1.79$, ns.) . In ROI analysis, we found an H x T interaction [$F(2, 54) = 4.43$, $p < .05$, $\eta^2 = .14$). As with midline analysis, frontal electrodes would appear with larger amplitudes than parietal. This effect was also more ample in the left hemisphere. Following the previous results in midline analysis, there were no differences between musicians and non-musicians [$F(1, 28) = 1.79$, ns.].

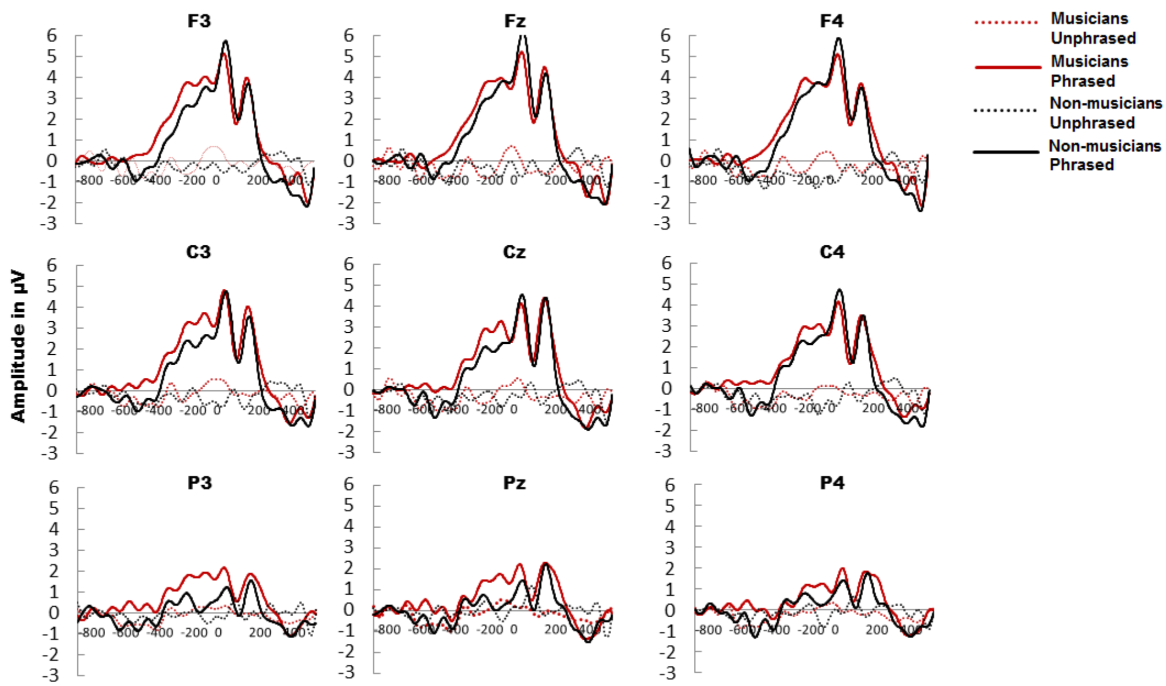

Figure 10. ERPs (averaged mean over frontal, central and parietal electrode sites) for Hummed condition for the time window -800 to 600 ms. Musicians in red, non-musicians in black. Dotted lines represent unphrased and solid lines phrased. Positive voltage is plotted up.

CPS in backwards analysis

In midline, we found a significant P x T interaction [$F(2, 54)=12.62$, $p < .001$, $\eta^2 = .32$]. Tukey post-hoc analysis showed that both frontal and central regions differed between phrased and unphrased conditions (*ps* < .001), while parietal regions were marginally non-significant ($p = .07$). In ROI analysis, we found an P x H x T interaction [$F(2, 54) = 3.58$, $p < .05$, $\eta^2 = .11$]. Further ANOVAs across topography revealed a main effect of phrasing in frontal electrodes [$F(1, 28) = 29.45$, $p < .001$, $\eta^2 = .51$]. In central electrodes, main effects of phrasing

[F(1, 28) = 30,19, p < .001, η^2 = .52] and Hemisphere [F(1, 28) = 16.93, p < .001, η^2 = .38] were found. Similarly, parietal regions also revealed main effects of phrasing [F(1, 28) = 7.74, p < .01, η^2 = .21] and hemisphere [F(1, 28) = 5.58, p < .05, η^2 = .17]. In regions of interest, we found a marginally significant main effect of group [F(1, 28) = 3.16, p = .086).

Source localization for CPS

We calculated swLORETA solution for grand averages of 200 to 700 ms for spoken sentences and 200 to 900 ms for hummed sentences. For presentation purposes, dipole activation magnitude and direction were overlaid with a T1 standard MRI brain scan, while the solution was carried in standard space. Dipoles are in respect to mean amplitudes of corresponding time-windows.

Spoken sentences

In spoken sentences we found two sources: a dipole on left superior frontal gyrus (Talairach coordinates: X = -39, Y = 48, Z = 26), a source in caudate body and head (Talairach coordinates: X = 2, Y = 10, Z = 7) and a source in the right superior temporal gyrus (Talairach coordinates: X = 54, Y = -4, Z = 5).

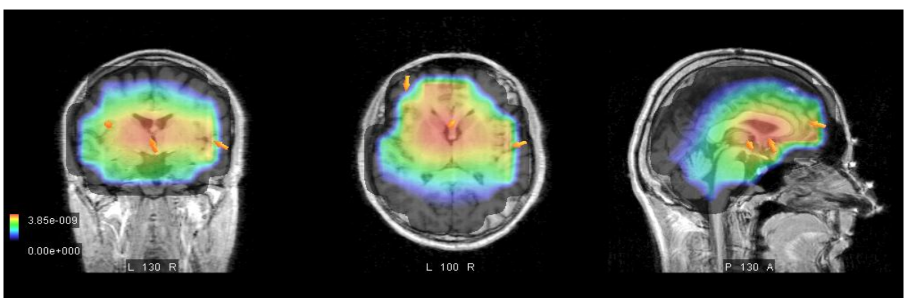

Figure 11. swLoreta analysis for source localization of CPS for spoken condition in musicians.

For non-musicians, we found two sources with significantly less activation magnitude. Those were found in left superior frontal gyrus (Talairach coordinates: X = -39, Y = 48, Z = 24) and in the right pre-central gyrus (Talairach coordinates: X = 60, Y = -5, Z = 19).

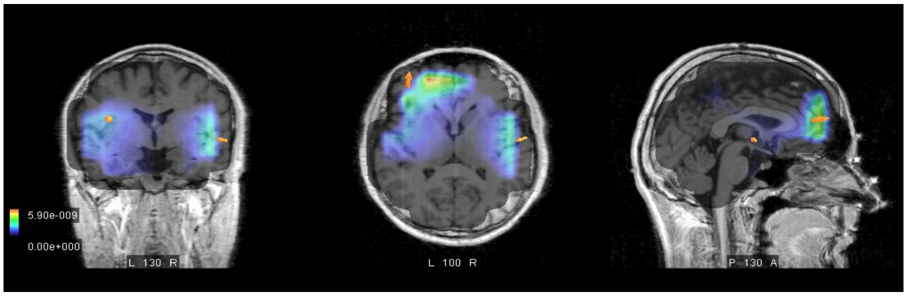

Figure 12. swLoreta analysis for source localization of CPS for spoken condition in non-musicians.

Hummed sentences

For hummed sentences, musicians showed two significant dipoles: the first in middle frontal gyrus (Talairach coordinates: X = -46, Y = 45, Z = 27) and the second in the inferior frontal gyrus (Talairach coordinates: X = -42, Y = 17, Z = -10).

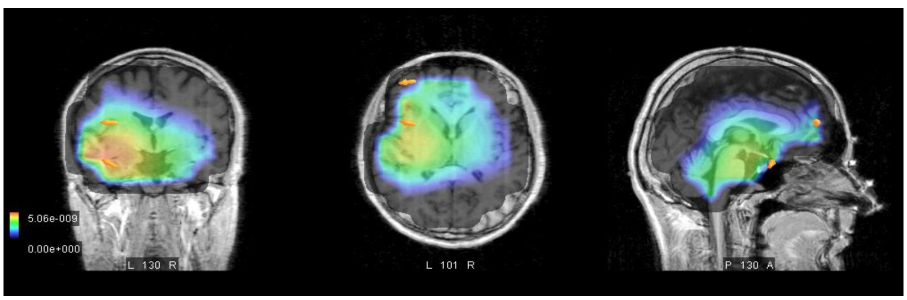

Figure 13. swLoreta analysis for source localization of CPS for spoken condition in musicians.

In non-musicians, three sources were found: the first in the superior frontal gyrus (Talairach coordinates: X = -38, Y = 49, Z = 26), the second in middle frontal gyrus (Talairach coordinates: X = -4, Y = 53, Z = 8), and third in the right para-hipocampal gyrus (Talairach coordinates: X = 23, Y = -17, Z = -16).

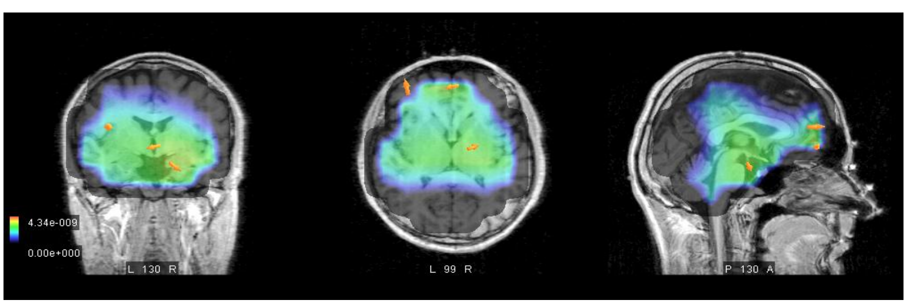

Figure 14. swLoreta analysis for source localization of CPS for spoken condition in non-musicians.

4. Discussion

Phrasing effects

We observed evidence of processing prosodic boundaries in the ERP traces. A positive component was detected after a prosodic boundary in the 100 to 1000 ms time window for spoken sentences and 200-1200 ms for hummed sentences. This phrasing effect was found in all mentioned time-windows when comparing phrased vs. unphrased sentences (the *F* values appear to be related with peak amplitude, cf. Figure 5 and Table 3). This finding accords well with Steinhauer et al. (1999), who observed for the first time that the end of an intonational phrase elicits a positive peak in ERPs. However, unlike Steinhauer et al. we found a fronto-central topography, not a fronto-parietal one. This may be a consequence of the method of analysis: whereas Steinhauer performed a whole-sentence analysis across all sentences, our was time-locked to the prosodic boundaries in each sentence. As far as we know, there have been only two studies that have aligned the phrasing onset across conditions with the end of the intonational phrase, Kerkhofs et al. (2008) and Mannel et al. (2011). When trigger points are aligned, the CPS is predominantly fronto-central (Mannel et al., 2011), whereas in whole-sentence analysis the topography appears in central and parietal regions. Due to the variation of the duration of pauses in whole-sentence analyses, we might argue that the time-locked method yields more reliable findings. Furthermore, we show that even when considering for the typically frontal exogenous components, the CPS appears more prominently in frontal regions.

Hummed vs. spoken comparison

Although the hummed and the spoken conditions are not strictly analogous due to the difference in experimental tasks, a comparison between them reveals that the lack of lexical content changes the topography of the CPS. In hummed sentences, the CPS was more anterior than in spoken sentences. A similar results has been reported by Pannekamp et al. (2005). We thus have converging evidence of stronger engagement of frontal resources in delexicalized, impoverished, speech. We also show that the CPS in hummed sentences, ending at 1200 ms, seems to last longer than the CPS in spoken, which returns to baseline level at 900 ms.

Additionally, by doing source localization we found common sources that seem to contribute to CPS. The regions close to the left superior temporal gyrus seem to play an important role in both hummed and spoken sentences. This goes in line with fMRI evidence showing the involvement of the superior temporal gyrus in phrasing (Ishebeck et al. 2008).

Closure positive shift and exogenous components

As is characteristic of research on phrasing, in our materials after a prosodic boundary, the next phrase started. Since the CPS is a component time-locked to the boundary, in one way or another, it is possible that the start of new information elicits a N100-P200 complex that could influence the CPS waveform. As stressed by Luck (2005), when components have variable onset latencies, latency jitter can occur therefore masking latent ERPs. Taking into account where the critical time windows of our results lie, it cannot be excluded that P50 and P200 add into the CPS amplitude or that N100 could mask the CPS.

Some studies indicate that N100-P200 complex does not play a decisive role in CPS. Mannel et al. (2011) excluded that such exogenous components could play a role in CPS by controlling the latency in which the N100-P200 complex would occur. Steinhauer (2003) argued that CPS could not be explained away by the P200 amplitude by showing that the CPS starts before the onset of the following phrase. Also, by applying a low-pass filter of 1 Hz, P200 effects were diminished while the CPS remained consistent (ib.). Such findings show that although P50 and P200 cannot be excluded from the observed positive peak, they play a small role in the CPS waveform. In this study, if we analyze the different time-windows we can observe that the CPS is significant in the 100-200 ms time-window for spoken sentences and again after 300 to 900 ms. The average pause duration in spoken sentences is 376±116 ms (cf. Table 1 in Methods). If the CPS could be explained by exogenous components, then in spoken sentences that effect could only be observed in the interval of those latencies, that is from 260 to 492 ms, plus the onset of P50 or P200; P50 amplitude would appear between 310-542 ms, while the P200 would appear between 460-692 ms. Also, the amplitude of N100, that would diminish the mean amplitude of the ERPs, would occur between 360 and 492 ms. In this line of reasoning, th CPS could only be explained by P50 amplitude if the peak started at least 260 ms after pause onset. And this is not what happens, the positivity starts before.

Another argument to consider is whether exogenous components might justify the differences in amplitude between musicians and non-musicians, since it has been reported that musicians have larger exogenous components than non-musicians (e.g. Marie et al., 2011). To exclude this alternative, we performed analyses on trigger 5, at the start of the new intonational phrase. We found no differences between musicians and non-musicians in any of the exogenous components. Thus we are confident that differences in amplitude between musicians and non-musicians are not attributable to these components. Furthermore, by doing a backwards analysis on trigger 5 we showed that the CPS is elicited around 500 ms before the onset of the next intonational phrase, that is, before the latency on which exogenous components appear. This is evidence that CPS cannot be explained away by exogenous components, even though small amplitude differences might contribute to the observed ERP traces, similarly in musicians and non-musicians.

Latency effects

Analyses of the latency of maximum amplitude revealed no significant statistical results. However, due to the pattern of CPS, maximum latency is not a good index of the latency effects. As seen in the ERPs, the differences between conditions appear predominantly in the onset of the CPS and not on the component peak. Since most of the ERP analyses on latency use peak data, we studied the differences in latency by analyzing the mean amplitude in early and latter time windows. We found differences between musicians and non-musicians in the onset and offset latency of CPS. In spoken sentences, interactions of group and phrasing were observed in early time-windows: the CPS has a shorter latency in musicians than non-musicians. This shows that musical expertise can influence the latency on which certain ERP components appear, as found studies in musical expertise in a variety of experimental paradigms (e.g., Marques et al., 2007; Nan et al. 2009).

In CPS, the starting latency has been under discussion, as there is no agreement on the onset latency of CPS and what factors underlie these differences. For example, Li and Yang (2009) have shown that the latency in which CPS appears seems to be dependent on the amount of information to process: the longer the sentence, the later the onset of the CPS. This points to a link between the onset of CPS and the effort required to process intonational phrases. We add to this evidence by showing that in musicians the CPS is elicited faster than in non-musicians. A likely reason is that the specialization of neural circuitry promoted by musical training would render phrasing less demanding cognitively, thus less effortful. The onset latency may depend on how the neural pathways are encoded to fulfill the roles of detecting acoustic cues to guide and integrate a sentence. An interesting effect is that this difference only seems to occur in spoken but not hummed sentences. If as argued in the OPERA hypothesis, the phrasing advantages in musicians comes from enhancement of low-level processing of acoustic cues like pitch and timing, then processing hummed sentences would be facilitated by prosodic cues and the absence of any lexical information. Inversely, speech should come closer between the two groups. A possible explanation for this result is the exigency of the experimental task. In fact, when listing to purely prosodic sentences, the attention required to segment information does not play such a strong role than when lexical information is present. Also, the task itself, being easier than the spoken counterpart, could have demanded less effort from the subjects. In line with the proposal of Knosche et al. (2005), where the CPS is a component that reflects the transition of attention between an intonational phrase to the next, it would be expected that this difference in latency would only occur when there is substantial information to parse. Our results go in line with this proposal, as musicians seem to have this advantage only when attention mechanisms are significatively elicited. Finally, spoken sentences require an examination of syntax rules, semantic meaning, and prosodic cues. As previously discussed, eight-month-old infants (pre-language acquisition) show a CPS, but only after 2 seconds, revealing a poorly developed mechanism to process phrasing (Pannekamp, 2006). Also, syntactic information seems essential to guide the phrasing of sentences, as CPS was only existent in children who developed syntactic rule knowledge (Mannel et al., 2011). The presence of syntax may therefore aid in phrasing processes. These differences in latency may then be explained by the combination of syntactic knowledge with the advantages of musical experts in processing prosodic information. This also shows how the enhancement of musicians in phrasing may not occur solely on a subcortical level as guided by automated pitch and timing decoding systems, but also by the interaction of these advantages with a more complex brain mechanism to interpret the acoustical cues.

Our results also show how the latency does not seem to depend on the duration of pauses, as questioned by Li & Yang (2010). To this date, no study showed a clear effect of latency in the CPS by comparing different conditions with similar pause durations. This raises the issue of whether this ERP component can be used to study other attributes of language parsing, by providing a direct measurement of brain mechanisms used to segment sentences.

This can be important to understand exactly how our brain develops to deal with the increasing demands, and cues, to segment information as discussed in Mannel et al.'s (2011) developmental study.

Amplitude effects

We found effects of musical expertise on the ERP amplitude for both spoken and hummed sentences. In spoken sentences, there were amplitude differences between musicians and non-musicians in the range 300-400 ms across the whole scalp and in the range 800-900 ms in frontal and central areas (several marginal differences in other time windows). In hummed sentences, differences were found in 100-200, 500-800 and 900-1000 ms time windows. Interestingly, both early and latter latency differences show main effects of group, showing overall larger amplitudes independently of the CPS. In hummed 500-700 time window, these differences show that musicians have larger CPS than non-musicians, but only in the left hemisphere. To our knowledge, no other studies revealed amplitude differences between musicians and non-musicians in the CPS using ERPs; they have only been reported by Neuhaus et al. (2006) in a study with musical phrases using MEG. These differences were interpreted as different processing strategies for musicians and non-musicians. Non-musicians would analyze the musical phrase boundary as an interruption, while musicians as a segmentation of a musical phrase as a whole. Musicians seem to focus on the prosodic contour, analyzing it in a holistic manner, while non-musicians focus on perceiving local small cues. This could be evidence of relying more or less strongly on holistic versus local processing strategies: non-musicians would rely more on local processing, whereas musicians would group into a Gestalt the sequences of successive phrases. Speculating on our data, the fact that musicians had higher CPS amplitudes for hummed than for spoken sentences might be due to their acoustic processing advantage. The absence of such differences in spoken sentences also contributes to this argument, since with the increment of lexical and syntactic cues, non-musicians can predict the end of intonational phrases as efficiently as musicians.

Laterality effects

As for CPS laterality, we found significant results between hemispheres. In spoken sentences, only the left hemisphere had phrasing effects in the 700-800 time window. In hummed sentences, several effects were found only in the left ROIs. The number of results in the left hemisphere shows how speech processing, commonly assigned to left brain areas, may be involved in phrasing. Interestingly, with source localization we also found that hummed sentences seem to have more proeminent left sources, but only in musicians, with the involvement of the left superior frontal gyrus and middle frontal gyrus. This fits in well with fMRI data showing that phrasing hummed sentences recruit more areas in the left hemisphere, specifically, the left supramarginal gyrus, the left superior frontal gyrus and the left Heschl gyrus (Ishebeck et al., 2008). Interestingly, we found Group effects in time-windows 500-700, but also only in the left hemisphere and in frontal regions, where musicians show higher amplitudes than non-musicians. The lateralization of group effects in hummed speech shows how the left hemisphere is not only important to hummed CPS, but also how musicians seem to have different neural mechanisms to process purely prosodical sentences.

5. Conclusion

The goal of the present study was to investigate whether musical expertise influences speech parsing. If musicians have improved neural pathways to process acoustic cues that are shared in music and language, then musical expertise might enhance the way in which different cues are weighted to group intonational phrases in a prosodic representation of a sentence. Our rationale was that as musicians excel in processing acoustical stimuli such as pitch and timing, as put forward by the OPERA hypothesis, then CPS would reflect how these processes may be facilitated. Since the CPS is a recently uncovered ERP component, many questions are still under debate. It was our goal to better understand the characteristics and functional role of this component.

Our findings contribute to clarify some of the open questions. The research on brain plasticity has grown exponentially in the last years, and while much was discovered about speech perception, few studies directly measured the contribution of these advantages on prosodic grouping of sentences with and without lexical content. Our results suggest that musical training enhances neural pathways that converge between lexical, syntactic and prosodic information, and how these advantages apply in phrasing processes; possibly in a more holistic way, in the case of musicians; or by analyzing local cues in the case of non-musicians. These differences in processing have various consequences: the most notable is the difference in latency, where musicians anticipate the perception of a phrase boundary with the combined aid of lexical and prosodic information. This shows an effect of expertise that enables musicians to use syntactic and prosodical cues efficiently. Differences in amplitude were found in hummed sentences, showing that musicians process them in a more holistic way than non-musicians, who focus on local cues to guide the closure of an intonational phrase. These results accor well with research on expertise suggesting that experts perceive, anticipate and process tasks in their domain of expertise in a more holistic, though differentiated, manner (Cellier et al., 1997). Our results clarify some aspects of the CPS. We add to the hypothesis of Knoshe et al. (2005) that the CPS is related to attention mechanisms to integrate different sentences, as our results indicate that the advantages of musicians in onset latency are only significant when attention is required. We also add evidence on how hummed sentences seem to rely on different mechanisms than spoken sentences, with more activation of frontal areas.

Interestingly, we found that hummed conditions are more left-oriented in the brain. Also, since musicians had higher CPS amplitudes in the left hemisphere for hummed speech, we can speculate whether the lack of lexical information may elicit different mechanisms to guide sentence phrasing. Whether this reflects how hard it is to parse a boundary without lexical information is something that should be analyzed in other studies. In line with the OPERA hypothesis (Patel, 2011), processing advantages brought about by musical expertise are related only to low-level neural networks, but also on the integration of these pathways with higher level brain regions, responsible for decoding temporal and prosodic cues. Our results fit in well by showing that musical training affects how we process phrasing in hummed and spoken sentences, and how the strategies to do so are different. Whether these differences relate to functional or anatomical differences is an open question and is outside the scope of this study. Further studies should focus on understanding what occurs on a functional level, by comparing between musicians and non-musicians responses to phrasing using imaging techniques.

In sum, we showed how musical training may benefit phrasing processes in the brain, and how they are guided by complex, subcortical and cortical mechanisms to deal with different types of stimuli. We gained insight on some of the CPS characteristics, such as the effects of musical expertise on latency, amplitude and laterality of CPS. Our study has some limitations. By focusing on the ecological validity of the stimuli, pause durations were not efficiently controlled which led to difficulties in the comparison of spoken effects in trigger 4 analysis. Due to the presence of exogenous components, additional analysis had to be done to assure the found results were not due to higher amplitudes of such components. For this reason, we cannot truly estimate the amplitude of CPS or the CPS waveform. Also, due to technical reasons, the EEG recordings were overall noisy, which led to the exclusion of three subjects (two musicians and one non-musician). With a bigger sample we could have reduced the signal-to-noise ratio, while adding statistical power to our results. The lack of a third controlled expertise group does not exclude the possibility that musicians had overall better results due to other variables, such as general intelligence or early cognitive engagement into a demanding task. In the future, studies should focus on clarifying what processes subtend the found results. Also, it would be interesting to study these differences in a foreign language and understand the contribution of syntactic knowledge and lexical information in the modulation of CPS.

References

Abrams, D. A., Bhatara, A., Ryali, S., Balaban, E., Levitin, D. J., & Menon, V. (2011). Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cerebral Cortex, 21*(7), 1507-1518.

Banse, R., & Scherer, K. R. (1996). Acoustical profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*(3), 614-636.

Besson, M., & Friederici, A. D. (1998). Language and music: a comparative view. *Music Perception, 16*(1), 1-9.

Castro, S. L., & Lima, C. F. (2010). Recognizing emotions in spoken language: A validated set of Portuguese sentences and pseudosentences for research on emotional prosody. *Behavior Research Methods, 42*(1), 74-81.

Catford, J. C. (2001). *A practical introduction to phonetics* (2nd ed.). Oxford: Oxford University Press.

Cellier, J., Eyrolle, H. & Marine, C. (1997). Expertise in dynamic environments. *Ergonomics 40*, 28–50.

Chapin, H., Jantzen, K., Kelso, J. A., Steinberg, F., & Large, E. (2010). Dynamic emotional and neural responses to music depend on performance expression and listener experience. *PLoS One, 5*(12), e13812.

Chi, M., Glaser, R., & Farr, M. (1988). *The nature of expertise*. Hillsdale: Lawrence Erlbaum Associates.

Darwin, C. (1974). *A origem do homem e a seleção sexual.* São Paulo: Hemus.

Detterman, D. K., & Ruthsatz, J. M. (1999). Toward a more comprehensive theory of exceptional abilities. *Journal for the Education of the Gifted, 22*, 148−158.

Ericsson, K. A., Nandagopal, K., & Roring, R. W. (2005). Giftedness viewed from the expert performance perspective. *Journal for the Education of the Gifted, 28*(3), 287−311.

Farrington-Darby, T., & Wilson, J. R. (2006). The nature of expertise: a review. *Applied Ergonomics, 37*(1), 17-32.

Frazier, L., Carlson, K., & Clifton, C., Jr. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Science, 10*(6), 244-249.

Fritz, J., Elhilali, M., & Shamma, S. (2005). Active listening: task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex. *Hearing Research, 206*(1-2), 159-176.

Gaser, C., & Schlaug, G. (2003a). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience, 23*(27), 9240-9245.

Gaser, C., & Schlaug, G. (2003b). Gray matter differences between musicians and nonmusicians. *Annals of the New York Academy of Sciences, 999,* 514-517.

Gomes, I., & Castro, S. L. (2003). Porlex, a lexical database in European Portuguese. *Psychologica, 32,* 91-108.

Han, Y., Yang, H., Lv, Y. T., Zhu, C. Z., He, Y., Tang, H. H.,... Dong, Q. (2009). Gray matter density and white matter integrity in pianists' brain: a combined structural and diffusion tensor MRI study. *Neuroscience Letters, 459*(1), 3-6.

Herdener, M., Esposito, F., di Salle, F., Boller, C., Hilti, C. C., Habermeyer, B., . . . Cattapan-Ludewig, K. (2010). Musical training induces functional plasticity in human hippocampus. *Journal of Neuroscience, 30*(4), 1377-1384.

Hwang, H., & Steinhauer, K. (2011). Phrase length matters: the interplay between implicit prosody and syntax in Korean "garden path" sentences. *Journal of Cognitive Neuroscience, 23*(11), 3555-3575.

Ischebeck, A. K., Friederici, A. D., & Alter, K. (2008). Processing prosodic boundaries in natural and hummed speech: an FMRI study. *Cerebral Cortex, 18(3)*, 541-552.

Isel, F., Alter, K., & Friederici, A. D. (2005). Influence of prosodic information on the processing of split particles: ERP evidence from spoken German. *Journal of Cognitive Neuroscience, 17*(1), 154-167.

Itzhak, I., Pauker, E., Drury, J. E., Baum, S. R., & Steinhauer, K. (2010). Event-related potentials show online influence of lexical biases on prosodic processing. *Neuroreport, 21*(1), 8-13.

Karni, A., Meyer, G., Jezzard, P., Adams, M. M., Turner, R., & Ungerleider, L. G. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature, 377*(6545), 155-158.

Kerkhofs, R., Vonk, W., Schriefers, H., & Chwilla, D. J. (2007). Discourse, syntax, and prosody: the brain reveals an immediate interaction. Journal of Cognitive *Neuroscience, 19*(9), 1421-1434.

Kerkhofs, R., Vonk, W., Schriefers, H., & Chwilla, D. J. (2008). Sentence processing in the visual and auditory modality: do comma and prosodic break have parallel functions? *Brain Research, 1224*, 102-118.

Knosche, T. R., Neuhaus, C., Haueisen, J., Alter, K., Maess, B., Witte, O. W., & Friederici, A. D. (2005). Perception of phrase structure in music. *Human Brain Mapping, 24*(4), 259-273.

Koelsch, S., Fritz, T., Schulze, K., Alsop D. & Schlaug, G. (2005). Adults and children processing music: An fMRI study. *Neuroimage, 25*, 1068–1076.

Kraus N, Skoe E, Parbery-Clark A, Ashley R. (2009) Experience-induced malleability in neural encoding of pitch, timbre and timing: implications for language and music. *Annals of the New York Academy of Sciences. 1169*, 543-557.

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience, 11*(8), 599-605.

Li, W., & Yang, Y. (2009). Perception of prosodic hierarchical boundaries in Mandarin Chinese sentences. *Neuroscience, 158*(4), 1416-1425.

Li, W., & Yang, Y. (2010). Perception of Chinese poem and its electrophysiological effects. *Neuroscience, 168*(3), 757-768.

Luck, S. (2005). *An introduction to the event-related potencial technique*. Cambridge : The Mit Press.

Magne, C., Astesano, C., Aramaki, M., Ystad, S., Kronland-Martinet, R., & Besson, M. (2007). Influence of syllabic lengthening on semantic processing in spoken French: behavioral and electrophysiological evidence. *Cerebral Cortex, 17*(11), 2659-2668.

Mannel, C., & Friederici, A. D. (2011). Intonational phrase structure processing at different stages of syntax acquisition: ERP studies in 2-, 3-, and 6-year-old children. Developmental Science, 14(4), 786-798.

Marie, C., Magne, C., & Besson, M. (2011). Musicians and the metric structure of words. *Journal of Cognitive Neuroscience, 23*(2), 294-305.

Marques, C., Moreno, S., Castro, S. L., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience, 19*(9), 1453-1463.

Moreno, S., & Besson, M. (2006). Musical training and language-related brain electrical activity in children. *Psychophysiology, 43*(3), 287-291.

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. *Cerebral Cortex*, *19*(3), 712-723.

Nan, Y., Knosche, T. R., & Friederici, A. D. (2006). The perception of musical phrase structure: a cross-cultural ERP study. *Brain Research, 1094*(1), 179-191.

Nan, Y., Knosche, T. R., & Friederici, A. D. (2009). Non-musicians' perception of phrase boundaries in music: A cross-cultural ERP study. *Biological Psychology, 82*(1), 70-81.

Neuhaus, C., Knosche, T. R., & Friederici, A. D. (2006). Effects of musical expertise and boundary markers on phrase perception in music. *Journal of Cognitive Neuroscience, 18*(3), 472-493.

Ohnishi, T., Matsuda, H., Asada, T., Aruga, M., Hirakata, M., Nishikawa, M., . . . Imabayashi, E. (2001). Functional anatomy of musical perception in musicians. *Cerebral Cortex, 11*(8), 754-760.

Oldfield, R. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia. 9*, 97-113.

Pannekamp, A., Toepel, U., Alter, K., Hahne, A., & Friederici, A. D. (2005). Prosody-driven sentence processing: an event-related brain potential study. *Journal of Cognitive Neuroscience, 17*(3), 407-421.

Pannekamp, A., Weber, C., & Friederici, A. D. (2006). Prosodic processing at the sentence level in infants. *Neuroreport, 17*(6), 675-678.

Pascual-Leone, A. (2001). The brain that plays music and is changed by it. *Annals of the New York Academy of Sciences, 930*, 315-329.

Pascual-Leone, A., Grafman, J., & Hallett, M. (1994). Modulation of cortical motor output maps during development of implicit and explicit knowledge. *Science, 263*(5151), 1287-1289.

Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA). Technical details: Methods Find. *Experimental Clinical Pharmacology, 24*(D), 5-12.

Patel, A. D. (2008). *Music, language, and the brain*. New York: Oxford University Press.

Patel, A. D. (2011). Why would Musical Training Benefit the Neural Encoding of Speech? The OPERA Hypothesis. *Frontiers in Psychology, 2,* 142.

Patel, A. D. (2012). The OPERA hypothesis: assumptions and clarifications. Annals of the New York Academy of Sciences, 1252, 124-128.

Patel, A.D., Peretz, I., Tramo, M. & Labrecque, R. (1998). Processing prosodic and musical patterns: A neuropsychological investigation. *Brain and Language, 61*(2), 123-144

Patel, A. D., & Iversen, J. R. (2007). The linguistic benefits of musical abilities. *Trends in Cognitive Sciences, 11*(9), 369-372.

Penhune, V. B. (2011). Sensitive periods in human development: evidence from musical training. *Cortex, 47*(9), 1126-1137.

Peretz, I., Cummings, S., & Dube, M. P. (2007). The genetics of congenital amusia (tone deafness): a family-aggregation study. *American Journal of Human Genetics, 81*(3), 582-588.

Pinker, S. (1999). How the mind works. *Annals of the New York Academy of Sciences, 882*, 119-127.

Polley, D. B., Steinberg, E. E., & Merzenich, M. M. (2006). Perceptual learning directs auditory cortical map reorganization through top-down influences. Journal of *Neuroscience, 26*(18), 4970-4982.

Roll, M., Lindgren, M., Alter, K., & Horne, M. (2012). Time-driven effects on parsing during reading. *Brain and Language, 121*(3), 267-272.

Rousseau, J.J. (1981). *Ensaio sobre a origem das línguas*. Lisboa: Estampa.

Ruthsatz, J., Detterman, D., Griscom, W. & Cirullo, B. (2008) Becoming an expert in the musical domain: It takes more than just practice. *Intelligence, 36*, 330–338.

Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience, 5*(7), 688-694.

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., Norton, A., Overy, K., & Winner, E. (2005). Effects of music training on the child's brain and cognitive development. *Annals of the New York Academy of Sciences, 1060*, 219-230.

Schon, K., Hasselmo, M. E., Lopresti, M. L., Tricarico, M. D., & Stern, C. E. (2004). Persistence of parahippocampal representation in the absence of stimulus input enhances long-term encoding: a functional magnetic resonance imaging study of subsequent memory after a delayed match-to-sample task. *Journal of Neuroscience, 24*(49), 11088-11097.

Schon, D., Magne, C., & Besson, M. (2004). The music of speech: music training facilitates pitch processing in both music and language. *Psychophysiology, 41*(3), 341-349.

Silva, S. (2005). *Parsing Speech and Parsing Music: Common Mechanisms*. Unpublished MA Thesis, Universidade de Porto, Portugal.

Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain and Language, 86*(1), 142-164.

Steinhauer, K., Abada, S. H., Pauker, E., Itzhak, I., & Baum, S. R. (2010). Prosody-syntax interactions in aging: event-related potentials reveal dissociations between on-line and off-line measures. *Neuroscience Letters, 472*(2), 133-138.

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience, 2*(2), 191-196.

Steinhauer, K., & Friederici, A. D. (2001). Prosodic boundaries, comma rules, and brain responses: the closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research, 30*(3), 267-295.

Toepel, U., Pannekamp, A., & Alter, K. (2007). Catching the news: Processing strategies in listening to dialogs as measured by ERPs. *Behavioral and Brain Functions, 3*, 53.

Vuust, P., Pallesen, K. J., Bailey, C., van Zuijen, T. L., Gjedde, A., Roepstorff, A., & Ostergaard, L. (2005). To musicians, the message is in the meter pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *Neuroimage, 24*(2), 560-564.

Wang, M. Q. & Hirschberg, J. (1992). Automatic classification of intonational phrasing boundaries. *Computer Speech and Language*, *6*(2), 175-196.

Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*(4), 420-422.

APPENDICES

APPENDIX I

Participant information

Laboratório de Fala FPCE-UP
labfala@fpce.up.pt

Rua do Dr. Manuel Pereira da Silva
4200-392 Porto, Portugal

Tel: +351 226 079 700
Fax: +351 226 079 725
Website: www.fpce.up.pt/labfala

Prof. São Luís F. Castro
Tel: +351 226 079 756
Email:slcastro@fpce.up.pt

Estudo sobre processamento prosódico
Informação ao Participante

Convidamos a participar neste estudo onde investigamos os processos implicados na percepção de fala fluente, em particular como organizamos a corrente acústica complexa que chega aos nossos ouvidos de modo a perceber as frases e os seus constituintes. Este é um estudo do Laboratório de Fala da FPCE-UP. Para que entenda o que a sua participação implica e por que razão este estudo é realizado, leia por favor a informação que se segue.

Quem pode participar no estudo?
Adultos com idades entre os 18 e os 45 anos.

O que envolve a participação?
A sua participação envolve a presença numa sessão experimental que dura aproximadamente 90 minutos, e que será conduzida por um psicólogo. A sessão realizar-se-á no Laboratório de Psicofisiologia da Faculdade. Será registada a actividade electroencefalográfica enquanto realiza uma tarefa simples de audição de frases. Para tal, ser-lhe á colocada uma touca com eléctrodos de modo a poder captar variações na actividade eléctrica em várias regiões do cérebro. As tarefas não provocam dor ou desconforto, nem têm riscos físicos ou psicológicos. As instruções serão devidamente explicadas e terá oportunidade de esclarecer dúvidas sempre que necessário.
Toda informação recolhida se destina exclusivamente a fins de investigação, e será tratada na mais estrita confidencialidade.

Sobre o que é o estudo?
Compreender sem esforço as frases faladas que ouvimos no dia-a-dia envolve um conjunto de operações mentais e cerebrais -- neurocognitivas -- que ainda hoje não conhecemos suficientemente bem, nem a ponto de ser capaz construir máquinas que reconheçam fala como um humano. Neste estudo examinaremos os recursos neurocognitivos que entram em funcionamento à medida que vai ouvindo frases, e como eles dependem do facto de ser possível ouvir frases propriamente ditas, ou ouvir frases aonde apenas está presente a “melodia da fala” (frases murmuradas, como se estivesse a cantarolar sem palavras). Os resultados contribuirão para compreender melhor como a comunicação através da linguagem é levada a cabo no nosso sistema mente/cérebro.

Quaisquer outros assuntos serão esclarecidos através de comunicação directa com os investigadores abaxo indicados. Agradecemos desde já a sua colaboração.

Paulo Branco: labfala@fpce.up.pt ; telefone: 917454744
Pedro Chaves: labfala@fpce.up.pt ; telefone: 220 400 610
Dr. Susana Silva: susanamsilva@fpce.up.pt; telefone: 939 544 210; 220 400 610
Prof. São Luís Castro: slcastro@fpce.up.pt ; telefone: 226 079 056

Apoios: Fundação para a Ciência e a Tecnologia

APPENDIX II

Informed consent form

Laboratório de Fala FPCE-UP
labfala@fpce.up.pt

Rua do Dr. Manuel Pereira da Silva
4200-392 Porto, Portugal

Tel: +351 226 079 700

Fax:+351 226 079 725
www.fpce.up.pt/labfala

Declaração de Consentimento

Esta declaração certifica que eu, ..., aceito de livre vontade participar na experiência científica Processamento prosódico A2, uma experiência que está integrada nas actividades de investigação do Laboratório de Fala desta Faculdade da Universidade do Porto, sob a supervisão da Prof. São Luís Castro.

Uma explicação breve sobre esta experiência e sobre o que a minha participação implica foi-me dada por .... Compreendi as explicações dadas, bem como os esclarecimentos que recebi a meu pedido.

A minha responsabilidade como participante é de participar activamente e com empenho. Se não estiver disponível para o fazer, exercerei o meu direito de desistir sem por isso ser penalizado/penalizada. No caso de decidir manter a minha participação, entendo que me comprometo a participar activamente.

Entendo também que caso deseje poderei vir a solicitar um sumário dos resultados do estudo.

....

Assinatura do Participante Data

Eu, abaixo assinado, dei as explicações sobre a investigação em causa.

....

Assinatura do Investigador Data

APPENDIX III

Instructions

Experiência Processamento Prosódico A2
Lab Fala

Instruções – Blocos Spoken

Variante SIM no botão Esquerdo

Vai ouvir várias frases faladas. Depois de cada frase terminar, verá uma palavra escrita. A tarefa que lhe pedimos é a de indicar se essa palavra escrita fazia parte, ou não fazia parte, da frase que acabou de ouvir. Se tiver ouvido essa palavra na frase, responda SIM. Se essa palavra não tiver estado na frase falada, responda, claro, Não.

Para responder **SIM**, carregue no **botão esquerdo** *[exemplificar com gesto]*; para responder **NAO**, carregue no **botão direito** *[exemplificar com gesto]*.

Um pouco antes de cada frase, existe um sinal visual – uma cruz branca.
Este sinal serve para avisar que logo a seguir vai ser apresentada a frase. Por isso, quando aparecer a cruz prepare-se para ouvir a frase.

Para que a recolha do EEG seja fidedigna, é importante não haver movimentos do olhar enquanto ouve a frase. Assim, enquanto ouve a frase por favor não faça movimentos com os olhos. Mantenha-os centrados nessa cruz.
Quando a cruz desaparecer, aproveite para piscar os olhos.

Variante SIM no botão Direito

Vai ouvir várias frases faladas. Depois de cada frase terminar, verá uma palavra escrita. A tarefa que lhe pedimos é a de indicar se essa palavra escrita fazia parte, ou não fazia parte, da frase que acabou de ouvir. Se tiver ouvido essa palavra na frase, responda SIM. Se essa palavra não tiver estado na frase falada, responda, claro, Não.

Para responder **SIM**, carregue no **botão direito** *[exemplificar com gesto]*; para responder **NAO**, carregue no **botão esquerdo** *[exemplificar com gesto]*.

Um pouco antes de cada frase, existe um sinal visual – uma cruz branca.
Este sinal serve para avisar que logo a seguir vai ser apresentada a frase. Por isso, quando aparecer a cruz prepare-se para ouvir a frase.

Para que a recolha do EEG seja fidedigna, é importante não haver movimentos do olhar enquanto ouve a frase. Assim, enquanto ouve a frase por favor não faça movimentos com os olhos. Mantenha-os centrados nessa cruz.
Quando a cruz desaparecer, aproveite para piscar os olhos.

Experiência Processamento Prosódico A2
Lab Fala

Instruções – Blocos Hummed

Variante SIM no botão Esquerdo

Vai ouvir várias frases ditas em “mmmm”, isto é, ditas com a boca fechada. Não são percebidas palavras, apenas a entoação. Em algumas destas frases, não em todas, existe também uma palavra falada.

Depois de cada frase terminar, verá uma palavra escrita. A tarefa que lhe pedimos é a de indicar se essa palavra escrita fazia parte, ou não fazia parte, da frase que acabou de ouvir. Se tiver ouvido essa palavra na frase, responda SIM. Se essa palavra não tiver estado na frase falada, responda, claro, Não.

Para responder **SIM**, carregue no **botão esquerdo** *[exemplificar com gesto]*;
para responder **NAO**, carregue no **botão direito** *[exemplificar com gesto]*.

Um pouco antes de cada frase, existe um sinal visual – uma cruz branca.
Este sinal serve para avisar que logo a seguir vai ser apresentada a frase. Por isso, quando aparecer a cruz prepare-se para ouvir a frase.

Para que a recolha do EEG seja fidedigna, é importante não haver movimentos do olhar enquanto ouve a frase. Assim, enquanto ouve a frase por favor não faça movimentos com os olhos. Mantenha-os centrados nessa cruz.
Quando a cruz desaparecer, aproveite para piscar os olhos.

Variante SIM no botão Direito

Vai ouvir várias frases ditas em “mmmm”, isto é, ditas com a boca fechada. Não são percebidas palavras, apenas a entoação. Em algumas destas frases, não em todas, existe também uma palavra falada.

Depois de cada frase terminar, verá uma palavra escrita. A tarefa que lhe pedimos é a de indicar se essa palavra escrita fazia parte, ou não fazia parte, da frase que acabou de ouvir. Se tiver ouvido essa palavra na frase, responda SIM. Se essa palavra não tiver estado na frase falada, responda, claro, Não.

Para responder **SIM**, carregue no **botão esquerdo** *[exemplificar com gesto]*;
para responder **NAO**, carregue no **botão direito** *[exemplificar com gesto]*.

Um pouco antes de cada frase, existe um sinal visual – uma cruz branca.
Este sinal serve para avisar que logo a seguir vai ser apresentada a frase. Por isso, quando aparecer a cruz prepare-se para ouvir a frase.

Para que a recolha do EEG seja fidedigna, é importante não haver movimentos do olhar enquanto ouve a frase. Assim, enquanto ouve a frase por favor não faça movimentos com os olhos. Mantenha-os centrados nessa cruz.
Quando a cruz desaparecer, aproveite para piscar os olhos.

APPENDIX IV

Participant questionnaire

Laboratório de Fala FPCE-UP
labfala@fpce.up.pt

Rua do Dr. Manuel Pereira da Silva
4200-392 Porto, Portugal

Tel: +351 226 079 700
Fax: +351 226 079 725
Website: www.fpce.up.pt/labfala

Prof. São Luís F. Castro
Tel: +351 226 079 756
Email:slcastro@fpce.up.pt

Questionário ao Participante

Processamento Prosódico A2

Nome:

Sexo:

Língua materna: Data

de nascimento:

Escolaridade (em anos):

Profissão:

N.º telemóvel/telefone: Email:

Informação sobre lateralidade

Que mão prefere para ... :	Sem Preferência	**Alguma vez usa a outra mão?**
Escrever	Esquerda ○ ○ ○ Direita	☐ Sim
Desenhar	Esquerda ○ ○ ○ Direita	☐ Sim
Atirar um objecto	Esquerda ○ ○ ○ Direita	☐ Sim
Usar a tesoura	Esquerda ○ ○ ○ Direita	☐ Sim
Usar a escova de dentes	Esquerda ○ ○ ○ Direita	☐ Sim
Usar a faca (sem garfo)	Esquerda ○ ○ ○ Direita	☐ Sim
Usar a colher	Esquerda ○ ○ ○ Direita	☐ Sim
Usar a vassoura (mão que fica em cima)	Esquerda ○ ○ ○ Direita	☐ Sim
Acender um fósforo	Esquerda ○ ○ ○ Direita	☐ Sim
Abrir uma caixa (mão na tampa)	Esquerda ○ ○ ○ Direita	☐ Sim
Usar o rato do computador	Esquerda ○ ○ ○ Direita	☐ Sim
Usar uma chave para abrir a porta	Esquerda ○ ○ ○ Direita	☐ Sim
Segurar num martelo	Esquerda ○ ○ ○ Direita	☐ Sim
Usar um pente ou escova	Esquerda ○ ○ ○ Direita	☐ Sim
Segurar numa chávena para beber	Esquerda ○ ○ ○ Direita	☐ Sim

Outras informações

Como avalia a sua acuidade auditiva? Muito Boa – 1 2 3 4 5 6 – Muito Má

Sofre actualmente de doenças neurológicas e/ou psiquiátricas? Sim / Não

Se sim, está a tomar medicação? Sim / Não

APPENDIX V

List of sentences

#	Lex	Boundaries	Sp or Hum	SENTENCES
1	1	1	s	O João comprou carne, o Jorge e a Luísa trouxeram saladas e bebidas.
2	1	2	s	O João comprou carnes, o Jorge trouxe salada, e a Luísa trouxe bebidas.
3	1	1	h	O João comprou carne, o Jorge e a Luísa trouxeram saladas e bebidas.
4	1	2	h	O João comprou carnes, o Jorge trouxe salada, e a Luísa trouxe bebidas.
5	2	1	s	A carne está estragada, o marisco e a fruta aguentaram-se bastante bem.
6	2	2	s	A carne está estragada, o marisco ficou bem, e a fruta aguentou-se firme.
7	2	1	h	A carne está estragada, o marisco e a fruta aguentaram-se bastante bem.
8	2	2	h	A carne está estragada, o marisco ficou bem, e a fruta aguentou-se firme.
9	3	1	s	O meu vestido era preto, o da Ana e o da Luísa tinham tons de azul e laranja.
10	3	2	s	O meu vestido era preto, o da Ana era todo azul, e o da Luísa tinha laranja.
11	3	1	h	O meu vestido era preto, o da Ana e o da Luísa tinham tons de azul e laranja.
12	3	2	h	O meu chapéu era preto, o da Ana era azul, e o da Luísa tinha laranja.
13	4	1	s	O Manuel apresentou, o Daniel e o Alexandre dançaram rumba e cha-cha-chá.
14	4	2	s	O Manuel apresentou, o Daniel dançou rumba, e o Alexandre cha-cha-chá.
15	4	1	h	O Manuel apresentou, o Daniel e o Alexandre dançaram rumba e cha-cha-chá.
16	4	2	h	O Zé Manel apresentou, a Sofia dançou rumba, e o Alexandre cha-cha-chá.
17	5	1	s	Eu fiquei por lá sentada, a Cláudia e a Inês sairam logo para a sala de jantar.
18	5	2	s	Eu fiquei por lá sentada, a Cláudia foi pr'ó átrio, e a Inês saiu pr'à cozinha.
19	5	1	h	Eu fiquei por lá sentada, a Cláudia e a Inês sairam logo para a sala de jantar.
20	5	2	h	Eu fiquei por lá sentada, a Cláudia foi pr'ó átrio, e a Inês saiu pr'à cozinha.
21	6	1	s	O avião é às três horas, o comboio da noite ou o barco já não interessam.
22	6	2	s	O avião é às três horas, o comboio sai às dez, e o barco já não interessa.
23	6	1	h	O avião é às três horas, o comboio da noite ou o barco já não interessam.
24	6	2	h	O avião é às três horas, o comboio sai às dez, e o barco já não interessa.
25	7	1	s	O exame é em Julho, as notas e o certificado só depois lá para Setembro.
26	7	2	s	O exame é em Julho, as notas em Agosto, e o certificado só em Outubro.
27	7	1	h	O exame é em Julho, as notas e o certificado só depois lá para Setembro.
28	7	2	h	O exame é em Julho, as notas em Agosto, e o certificado só em Setembro.
29	8	1	s	Perdi todo o subsídio, reclamei logo e contactei os serviços de finanças.
30	8	2	s	Perdi todo o subsídio, reclamei nas finanças, contactei a segurança social.
31	8	1	h	Perdi todo o subsídio, reclamei logo e contactei os serviços de finanças.
32	8	2	h	Perdi todo o subsídio, reclamei nas finanças, contactei a segurança social.
33	9	1	s	Eu saio de casa cedo, observo bem os ramos nus e os pássaros tão leves.
34	9	2	s	Eu saio de casa cedo, observo os ramos nus, e os pássaros tão leves.
35	9	1	h	Eu saio de casa cedo, observo bem os ramos nus e os pássaros tão leves.
36	9	2	h	Eu saio de casa cedo, observo os ramos nus, e os pássaros tão leves.
37	10	1	s	Nós limpamos a casa, os pais e você tratam da roupa e de abrir a porta.
38	10	2	s	Nós limpamos a casa, os pais tratam da roupa, e vocês abrem a porta.
39	10	1	h	Nós limpamos a casa, os pais e vocês tratam da roupa e de abrir a porta.
40	10	2	h	Nós limpamos a casa, os pais lavam a roupa, os primos arrumam a garagem.
41	11	1	s	Tu contas-me tudo já, eu e o juíz fazemos o relatório completo do caso.
42	11	2	s	Tu contas-me tudo já, eu faço o relatório, e depois o juíz vai expor o caso.
43	11	1	h	Tu contas-me tudo já, eu e o juíz fazemos o relatório completo do caso.

44	11	2	h	Tu contas-me tudo já, eu faço o relatório, e depois o juíz vai expor o caso.
45	12	1	s	Tu ensaias sozinha, o pianista e o encenador afinam-te e ensinam-te o gesto.
46	12	2	s	Tu ensaias sozinha, o pianista vai afinando, e o encenador ensina-te o gesto.
47	12	1	h	Tu ensaias sozinha, o pianista e o encenador afinam-te e ensinam-te o gesto.
48	12	2	h	Tu ensaias sozinha, o pianista vai afinando, e o encenador ensina-te o gesto.
49	13	1	s	Primeiro reflecti muito, depois acabei por consultar os melhores especialistas.
50	13	2	s	Primeiro reflecti muito, depois consultei técnicos, e ouvi vários especialistas.
51	13	1	h	Primeiro reflecti muito, depois acabei por consultar os melhores especialistas.
52	13	2	h	Primeiro reflecti muito, depois consultei técnicos, e ouvi vários especialistas.
53	14	1	s	Imprime-se tudo de manhã, à tarde o Daniel e a Joana separam as páginas.
54	14	2	s	Imprime-se tudo de manhã, à tarde o Rui separa, depois a Paula encaderna.
55	14	1	h	Imprime-se tudo de manhã, à tarde o Daniel e a Joana separam as páginas.
56	14	2	h	Imprime-se tudo de manhã, à tarde o Rui separa, depois a Paula encaderna.
57	15	1	s	Os cães dormem fora, os gatos e os peixes ficam onde estão dentro de casa.
58	15	2	s	Os cães dormem fora, os gatos ficam em casa, e os peixes estão no aquário.
59	15	1	h	Os cães dormem fora, os gatos e os peixes ficam onde estão dentro de casa.
60	15	2	h	Os cães dormem fora, os gatos ficam em casa, e os peixes estão no aquário.
61	16	1	s	Primeiro chegou tarde, depois pôs em questão e disse mal de tudo o que viu.
62	16	2	s	Primeiro chegou tarde, depois questionou o chefe, e no fim disse mal de tudo.
63	16	1	h	Primeiro chegou tarde, depois pôs em questão e disse mal de tudo o que viu.
64	16	2	h	Primeiro chegou tarde, depois questionou o chefe, e no fim disse mal de tudo.
65	17	1	s	O pai adorou a peça, mas a mãe e os filhos não gostaram mesmo nada.
66	17	2	s	O pai adorou a peça, mas a mãe detestou, e os filhos não gostaram nada.
67	17	1	h	O pai adorou a peça, mas a mãe e os filhos não gostaram mesmo nada.
68	17	2	h	O pai adorou a peça, mas a mãe detestou, e os filhos não gostaram nada.
69	18	1	s	Eu posso tratar de tudo, mas não tenho transporte para todos os instrumentos.
70	18	2	s	Eu posso tratar de tudo, mas não dos instrumentos, nem do respectivo transporte.
71	18	1	h	Eu posso tratar de tudo, mas não tenho transporte para todos os instrumentos.
72	18	2	h	Eu posso tratar de tudo, mas não dos instrumentos, nem do respectivo transporte.
73	19	1	s	Ele até ensina bem, mas os testes e os trabalhos são difíceis e numerosos.
74	19	2	s	Ele até ensina bem, mas os testes são longos, e os trabalhos muito difíceis.
75	19	1	h	Ele até ensina bem, mas os testes e os trabalhos são difíceis e numerosos.
76	19	2	h	Ele até ensina bem, mas os testes são longos, e os trabalhos muito difíceis.
77	20	1	s	Os cães são meigos, mas os gatos e então os peixes não têm nada disso.
78	20	2	s	Os cães são meigos, mas os gatos nada disso, e os peixes não interagem.
79	20	1	h	Os cães são meigos, mas os gatos e então os peixes não têm nada disso.
80	20	2	h	Os cães são meigos, mas os gatos nada disso, e os peixes não interagem.
81	21	1	s	Conferi várias vezes, mas mesmo assim só depois de reler é que enviei tudo.
82	21	2	s	Conferi várias vezes, mesmo assim ele releu, e só depois é que enviou tudo.
83	21	1	h	Conferi várias vezes, mas mesmo assim só depois de reler é que enviei tudo.
84	21	2	h	Conferi várias vezes, mesmo assim ele releu, e só depois é que enviou tudo.
85	22	1	s	Eu gosto bastante dela, mas a paciência e a compreensão por vezes falham.
86	22	2	s	Eu gosto bastante dela, mas a paciência falha, e compreensão é bem difícil.
87	22	1	h	Eu gosto bastante dela, mas a paciência e a compreensão por vezes falham.
88	22	2	h	Eu gosto bastante dela, mas a paciência falha, e compreensão é bem difícil.
89	23	1	s	O João trabalhou bem, mas o júri e o público não apreciaram nada o estilo.

90	23	2	s	O João trabalhou bem, mas o júri é soberano, e até o público desaprovou.
91	23	1	h	O João trabalhou bem, mas o júri e o público não apreciaram nada o estilo.
92	23	2	h	O João trabalhou bem, mas o júri é soberano, e até o público desaprovou.
93	24	1	s	Tornou-se conhecido, e deixou de se interessar como antes pelo rigor.
94	24	2	s	Tornou-se conhecido, mas ao ficar célebre, descurou o trabalho e o rigor.
95	24	1	h	Tornou-se conhecido, e deixou de se interessar como antes pelo rigor.
96	24	2	h	Tornou-se conhecido, mas ao ficar célebre, descurou o trabalho e o rigor.
97	25	1	s	A mesa estava bonita, mas a carne e as batatas estavam mal cozinhadas.
98	25	2	s	A mesa estava bonita, mas a carne era má, e as batatas estavam cruas.
99	25	1	h	A mesa estava bonita, mas a carne e as batatas estavam mal cozinhadas.
100	25	2	h	A mesa estava bonita, mas a carne era má, e as batatas estavam cruas.
101	26	1	s	Aspirei bem os quartos, mas mesmo assim o pó continuou no ar abafado.
102	26	2	s	Aspirei bem os quartos, mas o pó não saiu todo, nem o ar ficou fresco.
103	26	1	h	Aspirei bem os quartos, mas mesmo assim o pó continuou no ar abafado.
104	26	2	h	Aspirei bem os quartos, mas o pó não saiu todo, nem o ar ficou fresco.
105	27	1	s	Guardei tudo no armário, mas as malas e cobertas são grandes e pesadas.
106	27	2	s	Guardei tudo no armário, mas as malas não cabem, e as cobertas são pesadas.
107	27	1	h	Guardei tudo no armário, mas as malas e cobertas são grandes e pesadas.
108	27	2	h	Arrumei bem o armário, mas a mala não cabe, e as cobertas também não.
109	28	1	s	Ela quase que caía, mas o André e a Maria agarraram-na bem pelo braço.
110	28	2	s	Ela quase que caía, mas o André agarrou-a, e a Maria segurou-lhe o braço.
111	28	1	h	Ela quase que caía, mas o André e a Maria agarraram-na bem pelo braço.
112	28	2	h	Ela quase que caía, mas o André agarrou-a, e a Maria segurou-lhe o braço.
113	29	1	s	Eu preencho os papéis, mas ele e o sócio têm de reler e assinar depressa.
114	29	2	s	Eu preencho os papéis, mas ele tem de reler, e o sócio tem que assinar já.
115	29	1	h	Eu preencho os papéis, mas ele e o sócio têm de reler e assinar depressa.
116	29	2	h	Eu preencho os papéis, mas ele tem de reler, e o sócio tem que assinar já.
117	30	1	s	A mesa já é velha, mas a madeira e a cor são muito bonitas e requintadas.
118	30	2	s	A mesa já é velha, mas a madeira é boa, e a cor parece-me requintada.
119	30	1	h	A mesa já é velha, mas a madeira e a cor são muito bonitas e requintadas.
120	30	2	h	A mesa já é velha, mas a madeira é boa, e a cor parece-me requintada.
121	31	1	s	A Maria não cedeu, e apesar da insistência levou em frente a sua ideia.
122	31	2	s	A Maria não cedeu, apesar da insistência, e levou a sua ideia ávante.
123	31	1	h	A Maria não cedeu, e apesar da insistência levou em frente a sua ideia.
124	31	2	h	A Maria não cedeu, apesar da insistência, e levou a sua ideia ávante.
125	32	1	s	Embora esteja frio, já se sente um calorzinho do sol e um ar leve de verão.
126	32	2	s	Embora esteja frio, o sol já está brilhante, e sente-se um ar leve de verão.
127	32	1	h	Embora esteja frio, já se sente um calorzinho do sol e um ar leve de verão.
128	32	2	h	Embora esteja frio, o sol já está brilhante, e sente-se um ar leve de verão.
129	33	1	s	Quando forem horas, tu e o secretário tratam desses papéis e dos telefonemas.
130	33	2	s	Quando forem horas, tu tratas desses papéis, e o secretário faz os telefonemas.
131	33	1	h	Quando forem horas, tu e o secretário tratam desses papéis e dos telefonemas.
132	33	2	h	Quando forem horas, tu tratas desses papéis, e o secretário faz os telefonemas.
133	34	1	s	Segundo o que dizem, mãe e filha percebem muito de festas e recepções.
134	34	2	s	Segundo o que têm dito, a mãe percebe de festas, e a filha sabe receber.
135	34	1	h	Segundo o que dizem, mãe e filha percebem muito de festas e recepções.

136	34	2	h	Segundo o que dizem, a mãe percebe de festas, e a filha sabe receber.
137	35	1	s	Se nos fores lá buscar, a Isabel e eu levamos as duas colunas e o amplificador.
138	35	2	s	Se nos fores lá buscar, levo as duas colunas, e a Isabel traz o amplificador.
139	35	1	h	Se nos fores lá buscar, a Isabel e eu levamos as duas colunas e o amplificador.
140	35	2	h	Se nos fores lá buscar, levo as duas colunas, e a Isabel traz o amplificador.
141	36	1	s	No caso de chover, a câmara e a escola oferecem capas e guarda-chuvas.
142	36	2	s	No caso de chover, a câmara cede capas, e a escola dá guarda-chuvas.
143	36	1	h	No caso de chover, a câmara e a escola oferecem capas e guarda-chuvas.
144	36	2	h	No caso de chover, a câmara cede capas, e a escola dá guarda-chuvas.
145	37	1	s	Sempre que me encontra, tem o hábito de dizer piadas antes da conversa.
146	37	2	s	Sempre que me encontra, diz piadas e anedotas, e depois conversa a sério.
147	37	1	h	Sempre que me encontra, tem o hábito de dizer piadas antes da conversa.
148	37	2	h	Sempre que me encontra, diz piadas e anedotas, e depois conversa a sério.
149	38	1	s	Desde que ali entrou, deixou de se ouvir o barulho e a confusão de antes.
150	38	2	s	Desde que ali entrou, não se ouve barulho, nem houve mais confusão.
151	38	1	h	Desde que ali entrou, deixou de se ouvir o barulho e a confusão de antes.
152	38	2	h	Desde que ali entrou, não se ouve barulho, nem houve mais confusão.
153	39	1	s	Quando logo saires, não te esqueças de levar a chave e os teus postais.
154	39	2	s	Quando logo saires, fecha bem à chave, e leva embora os teus postais.
155	39	1	h	Quando logo saires, não te esqueças de levar a chave e os teus postais.
156	39	2	h	Quando logo saires, fecha bem à chave, e leva embora os teus postais.
157	40	1	s	Sempre que eu posso, gosto de passear a pé e andar pelas ruas da cidade.
158	40	2	s	Sempre que posso, dou passeios a pé, e vagueio pelas ruas da cidade.
159	40	1	h	Sempre que eu posso, gosto de passear a pé e andar pelas ruas da cidade.
160	40	2	h	Sempre que eu posso, dou passeios a pé, e vagueio pelas ruas da cidade.
161	41	1	s	Quando vocês chegarem, o João e eu vamos buscar-vos com a bagagem também.
162	41	2	s	Quando vocês chegarem, eu vou buscar-vos, e o João ajuda com a bagagem.
163	41	1	h	Quando vocês chegarem, o João e eu vamos buscar-vos com a bagagem também.
164	41	2	h	Quando vocês chegarem, eu vou buscar-vos, e o João ajuda com a bagagem.
165	42	1	s	Se houver desacordo, reunimos todos e acertamos qual é a melhor estratégia.
166	42	2	s	Se houver desacordo, reunimos as equipas, e acertamos a melhor estratégia.
167	42	1	h	Se houver desacordo, reunimos todos e acertamos qual é a melhor estratégia.
168	42	2	h	Se houver desacordo, reunimos as equipas, e acertamos a melhor estratégia.
169	43	1	s	Contando que haja sol, a Helena e o Carlos trazem a prancha e a mota de água.
170	43	2	s	Contando que haja sol, a Helena traz a prancha, e o Carlos a mota de água.
171	43	1	h	Contando que haja sol, a Helena e o Carlos trazem a prancha e a mota de água.
172	43	2	h	Contando que haja sol, a Helena traz a prancha, e o Carlos a mota de água.
173	44	1	s	Mal o carro apareceu, vítima e agressor foram agarrados e levados para dentro.
174	44	2	s	Mal o carro apareceu, a vítima foi agarrada, e o agressor foi levado para dentro.
175	44	1	h	Mal o carro apareceu, vítima e agressor foram agarrados e levados para dentro.
176	44	2	h	Mal o carro apareceu, a vítima foi agarrada, e o agressor foi levado para dentro.
177	45	1	s	Concordei com tudo, desde que pudesse ver e também experimentar por um dia.
178	45	2	s	Concordei com tudo, desde que pudesse ver, e depois experimentar por um dia.
179	45	1	h	Concordei com tudo, desde que pudesse ver e também experimentar por um dia.
180	45	2	h	Concordei com tudo, desde que pudesse ver, e depois experimentar por um dia.
181	46	1	s	Trabalhamos nesta sala, só se a Eva e o Pedro a pintarem e decorarem.

182	46	2	s	Trabalhamos nesta sala, se a Eva a pintar, e também se o Pedro a decorar.
183	46	1	h	Trabalhamos nesta sala, só se a Eva e o Pedro a pintarem e decorarem.
184	46	2	h	Trabalhamos nesta sala, se a Eva a pintar, e também se o Pedro a decorar.
185	47	1	s	Se for mesmo preciso, posso acabar ainda hoje as reportagens e as entrevistas.
186	47	2	s	Se for mesmo preciso, posso acabar isto hoje, e deixo para amanhã a entrevista.
187	47	1	h	Se for mesmo preciso, posso acabar ainda hoje as reportagens e as entrevistas.
188	47	2	h	Se for mesmo preciso, posso acabar isto hoje, e deixo para amanhã a entrevista.
189	48	1	s	Em situações destas, é melhor manter silêncio e também muita discrição.
190	48	2	s	Em situações destas, é melhor criar silêncio, e manter muita discrição.
191	48	1	h	Em situações destas, é melhor manter silêncio e também muita discrição.
192	48	2	h	Em situações destas, é melhor criar silêncio, e manter muita discrição.