Brain & Language 147 (2015) 51–57

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Short Communication

Differential oscillatory encoding of foreign speech

Alejandro Pérez [a,*], Manuel Carreiras [a,b,c], Margaret Gillon Dowens [d], Jon Andoni Duñabeitia [a]

[a] BCBL – Basque Center on Cognition Brain and Language, 20009 Donostia, Spain
[b] Ikerbasque, Basque Foundation for Science, 48011 Bilbao, Spain
[c] Departamento de Filología Vasca, EHU/UPV, 48015 Bilbao, Spain
[d] University of Nottingham Ningbo Interdisciplinary Centre on Research in Neuroscience (UNNICORN), 315100 Ningbo, China

ARTICLE INFO

Article history:
Received 6 September 2014
Accepted 16 May 2015
Available online 9 June 2015

Keywords:
Language
Oscillations
Bilingualism
Speech
Foreign language
Synchrony

ABSTRACT

Neuronal oscillations play a key role in auditory perception of verbal input, with the oscillatory rhythms of the brain showing synchronization with specific frequencies of speech. Here we investigated the neural oscillatory patterns associated with perceiving native, foreign, and unknown speech. Spectral power and phase synchronization were compared to those of a silent context. Power synchronization to native speech was found in frequency ranges corresponding to the theta band, while no synchronization patterns were found for the foreign speech context and the unknown language context. For phase synchrony, the native and unknown languages showed higher synchronization in the theta-band than the foreign language when compared to the silent condition. These results suggest that neural synchronization patterns are markedly different for native and foreign languages.

1. Introduction

Theories based on the correspondence between periods of different oscillatory rhythms in the brain (Buzsáki, 2006) and the periodic temporal fluctuations (quasi-rhythm) of the acoustic signal of speech (Ahissar & Ahissar, 2005; Peelle & Davis, 2012; Poeppel, 2003) suggest that neural oscillations play a key role in speech perception. Several theoretical 'speech-brain syncing' models specifically claim that intelligibility of speech is achieved when the cortical oscillations in theta, beta and gamma frequency bands remain phase-locked to the rhythm of the auditory input (Ghitza, 2011; Ghitza & Greenberg, 2009; Giraud & Poeppel, 2012; Luo & Poeppel, 2007). In particular, the neural 'theta oscillator' (4–10 Hz) has been proposed to lead speech perception by tracking the syllabic quasi-rhythm of the input and hierarchically setting the frequency of the beta and gamma oscillators (Ghitza, 2013; Ghitza, Giraud, & Poeppel, 2012).

Oscillatory analyses of nonnative speech perception within the speech-brain syncing framework suggest that neural synchronization in native and nonnative language contexts may be different (Jin, Diaz, Colomer, & Sebastian-Galles, 2014). Given the naturalness of conversational speech, in which phonetic, syntactic, semantic and pragmatic aspects of language are integrated (Reiterer, Pereda, & Bhattacharya, 2011), paradigms in which neural synchronization to well-formed longer samples of native vs. nonnative speech are compared to each other are clearly needed. So far, however, in spite of their obvious ecological validity, studies of brain oscillations during perception of long samples of nonnative speech such as dialogues are very scarce (Peña & Melloni, 2012; Reiterer et al., 2011). The current study was designed to address this issue by exploring the neural oscillatory patterns associated with nonnative foreign speech.

In this electroencephalographic (EEG) study, we assessed differences in the neural oscillatory patterns associated with processing continuous, close-to-natural speech in different language contexts: native speech, known foreign speech and unknown speech (which were compared to a silent condition). An experimental scenario was created where native Spanish speakers who were relatively proficient in English (i.e., foreign language) but had no knowledge of French (i.e., unknown language) completed a visual go/no-go task while different audio tracks including everyday conversations were being played in the background (a manipulation that was completely orthogonal to the main visual discrimination task). Because amplitude (power) and phase synchronization can reflect different neural processes (i.e., Palva & Palva, 2007), both were investigated here. Spectral power reflects local activations of large groups of neurons (Singer, 1999), indexing the resonance of the brain activity to the rhythms of speech (Buiatti, Peña, & Dehaene-Lambertz, 2009). In contrast, phase coupling between different EEG signals (e.g., pairs of electrodes) is typically considered

* Corresponding author at: Paseo Mikeletegi 69, 2, 20009 Donostia, Spain. Fax: +34 943 309 052.
E-mail address: a.perez@bcbl.eu (A. Pérez).

http://dx.doi.org/10.1016/j.bandl.2015.05.008
0093-934X/

an indication of the functional relationship between (distant) brain sites (see Sauseng & Klimesch, 2008). Phase-locking measures can therefore be used to explore long-range interactions between brain regions (Varela, Lachaux, Rodriguez, & Martinerie, 2001), and are especially relevant since neural synchronization to speech has been shown to be the result of phase resetting of ongoing neural oscillations (Luo & Poeppel, 2007). In the current study, the phase coupling for investigating brain functional relationships was assessed through the phase locking value (PLV) measure, which is an amplitude independent measure of the relative constancy of the difference of phases between two signals (Lachaux, Rodriguez, Martinerie, & Varela, 1999), here the entire EEG signal associated with processing conversational speech.

Listening to native speech was expected to yield local and global synchrony of specific neural assemblies, mainly in the theta frequency band (Howard & Poeppel, 2012). It was also hypothesized that listening to an unknown language would elicit significant (yet reduced) neural theta-band synchronization, in line with demonstrations of stimulus-locked theta synchronization for unintelligible and unknown speech (Howard & Poeppel, 2010; Peña & Melloni, 2012). In the same line, for the known foreign language, an oscillatory neural pattern similar to that of the native language could initially be expected, although non-balanced bilinguals are unlikely to decode and comprehend foreign speech to the same extent as native speech.

Interestingly, recent evidence showed that processing of nonnative (as compared to native) sub-lexical units results in a different pattern of neural synchronization in the theta-band (Jin et al., 2014). In addition, in native language contexts under circumstances in which language processing has relatively high cognitive costs (e.g. for Dyslexic readers), a reduced neural synchronization may be observed (Soltész, Szűcs, Leong, White, & Goswami, 2013). Since foreign speech processing also involves high cognitive cost (Perani & Abutalebi, 2005), it could be tentatively predicted to show reduced synchronization as compared to native speech. The results of this study will contribute to a better definition of the neural mechanisms involved in processing of native and foreign speech.

2. Methods

2.1. Participants

Twenty-four right-handed native Spanish speakers (mean age = 22.96 years, SD = 3.92, 9 females) took part in the study. None of them had any prior knowledge of French but reported a mean self-perceived English proficiency of 6.42 (SD = 1.1; 1-to-10 scale) and an average age of acquisition of English of 6.75 years (SD = 2.8). Participants' linguistic abilities in English were rated (following a 1-to-5 scale) in an individual interview (mean = 3.38, SD = 0.65). Furthermore, participants completed a Spanish and English picture-naming test including 77 drawings of common objects. Results in the Spanish version showed a close-to-perfect performance (mean = 76.63, SD = 0.71) while results in the English version showed a relatively good nonnative vocabulary level (mean = 55.29, SD = 10.09). Finally, an English lexical decision task was administered, in which participants were asked to decide whether visually displayed strings corresponded to an existing English word ($N = 50$) or not ($N = 50$). Results showed very low error rates (mean = 2.58%, SD = 2.60), indicating that participants were perfectly able to discriminate between existing English words and nonwords. All participants had normal visual and hearing acuity and no past history of psychiatric or neurological disorders. The study was carried out in accordance with the principles laid down in the Declaration of Helsinki, and approved by the BCBL Ethics Committee.

2.2. Materials and procedure

2.2.1. Visual go/no-go task

The experimental task consisted of a visual go/no-go task. Each trial started with a black screen (5000 ms), after which a star appeared in the center of the screen. Participants were instructed to press a button if the star t was red (which occurred in 10% of the trials) and to withhold their response if it was white. The star remained on the screen for 2500 ms, or until a response was collected. A total of 320 trials were presented across four identical 10-min blocks (80 trials per block) separated by short breaks.

2.2.2. Background audio

Critically, blocks were presented under different auditory conditions. One block was completed in a silent environment (No Audio condition), whereas the other blocks included the presentation of a 10-min audio track played in the background through loudspeakers. The audio tracks were taken from the same scene from the film 'Kramer vs. Kramer', that included extended dialogues among different characters. In one block, the audio track corresponded to the Spanish version of the film (Native Language). In another block, the audio was English (Foreign Language), and the other block included the same scene in French (Unknown Language). There were no significant differences in spectral power between the different speech streams. Importantly, participants were not informed about the presence or nature of the speech (i.e. the audio tracks), and they were not explicitly told to pay attention to it. All possible combinations of block order presentation ($N = 24$) were administered.

2.2.3. Recall test

Once the four blocks had been completed, participants were asked to complete a recall test in which they had to decide whether a series of visually-presented short sentences were part of the speech streams that they had heard during the visual go/no-go task. The task included 80 different sentences (40 Spanish, 40 English) from the same film, of which 50% was actually presented during the experiment (20 sentences per language). The sentences were visually presented in the center of the screen in a random order. Participants were asked to press one button if they thought that they heard it before and a different button if they thought that the sentence had not been presented. Sentences were matched for length, had a similar syntactic structure, and were selected so that successful completion of the recall test necessarily required direct understanding of the soundtrack (e.g., English present: "They set the court date", "I got myself a job", "It's still a workday"; English absent: "We listened to the radio", "He is in first grade", "We need detergent").

2.3. EEG recording and analysis

Participants were seated comfortably in a sound-attenuated Faraday chamber with a monitor positioned 80–90 cm in front of them (at eye-level). They were instructed to remain quiet and to avoid body movement. Electrophysiological data was acquired using an elastic cap with a 32-channel BrainAmp system (Brain Products GmbH). Left mastoid was used as reference. Eye movements and blinks were monitored with four additional electrodes providing recordings of the horizontal and vertical EOG. Inter-electrode impedances were kept below 10 kΩ. Data were acquired at a sampling rate of 250 Hz. Using Brain Vision Analyzer (version 2.0.2), the EEG signal was off-line re-referenced to the averaged activity of the two mastoids. A high-pass filter of 0.1 Hz was applied before performing an ICA. Identified components were visually inspected searching for those accounting for ocular movements in order to remove them from the data ($\leqslant 4$