

Accepted Manuscript

Differential brain-to-brain entrainment while speaking and listening in native and foreign languages

Alejandro Pérez, Guillaume Dumas, Melek Karadag, Jon Andoni Duñabeitia



PII: S0010-9452(18)30405-2

DOI: <https://doi.org/10.1016/j.cortex.2018.11.026>

Reference: CORTEX 2489

To appear in: *Cortex*

Received Date: 3 July 2018

Revised Date: 28 September 2018

Accepted Date: 29 November 2018

Please cite this article as: Pérez A, Dumas G, Karadag M, Duñabeitia JA, Differential brain-to-brain entrainment while speaking and listening in native and foreign languages, *CORTEX*, <https://doi.org/10.1016/j.cortex.2018.11.026>.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Title: Differential brain-to-brain entrainment while speaking and listening in native and foreign languages

Author Names: Alejandro Pérez^{1,2,3*}, Guillaume Dumas^{4,5,6}, Melek Karadag⁷ & Jon Andoni Duñabeitia^{3,8}

Author Affiliations:

¹ Centre for French & Linguistics, University of Toronto Scarborough, M1C 1A4, Toronto, Canada.

² Psychology Department, University of Toronto Scarborough, M1C 1A4, Toronto, Canada.

³ BCBL, Basque Center on Cognition Brain and Language, 20009, Donostia-San Sebastián, Spain.

⁴ Human Genetics and Cognitive Functions Unit, Institut Pasteur, 75015 Paris, France.

⁵ CNRS UMR 3571 Genes, Synapses and Cognition, Institut Pasteur, 75015 Paris, France.

⁶ Human Genetics and Cognitive Functions, University Paris Diderot, Sorbonne Paris Cité, 75013 Paris, France.

⁷ Centre for Speech, Language and the Brain, Department of Psychology, University of Cambridge, CB2 3EB, Cambridge, United Kingdom.

⁸ Facultad de Lenguas y Educación, Universidad Nebrija, 28015, Madrid, Spain.

*Corresponding author.

Corresponding author: Alejandro Pérez, e-mail: alejandro.perez@utoronto.ca

Postal address: 1265 Military Trail, Toronto, ON, Canada, M1C 1A4

Telephone: +1 647 871 5713

Declarations of interest: none

1. Introduction

Verbal communication is associated with synchronous patterns of brain activity between the speaker and the listener (Dai et al., 2018; Jiang et al., 2012; Kawasaki, Yamada, Ushiku, Miyauchi, & Yamaguchi, 2013; Kuhlen, Allefeld, & Haynes, 2012; Liu et al., 2017; Nozawa, Sasaki, Sakaki, Yokoyama, & Kawashima, 2016; Spiegelhalder et al., 2014; Tadic, Andjelkovic, Boshkoska, & Levnajic, 2016). This interbrain coupling seems to reflect the alignment between the processes of speech production and comprehension necessary for mutual understanding and successful communication (Dikker, Silbert, Hasson, & Zevin, 2014; Schoot, Hagoort, & Segaert, 2016; Stephens, Silbert, & Hasson, 2010). Interestingly, in spite of the crucial importance of face-to-face interactions (Jiang et al., 2012), interbrain coupling also takes place in a turn-taking verbal exchange under conditions of no interpersonal visual contact (Ahn et al., 2017; Pérez, Carreiras, & Duñabeitia, 2017). In other words, successful verbal interaction is associated with brain-to-brain synchronization regardless of non-verbal communication. This indicates that, at least in healthy adults using their native language, a shared speech signal and the willingness to communicate are enough to elicit interbrain synchronization. This phenomenon has been termed brain-to-brain entrainment during speaking and listening (henceforth, B2B), and it has been suggested to result from the mutual neural entrainment mediated by the speech signal (Pérez et al., 2017).

Neural entrainment is a general mechanism by which brain oscillations adapt to the rhythm of regular stimulation (Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008). This way, and as a consequence of a series of predictive mechanisms, the upcoming stimuli in a time series will coincide with a specific predicted phase of the entrained oscillation, leading to an amplification of the neural response to expected stimulus and attenuation of those not expected (Lakatos, Chen, O'Connell, Mills, & Schroeder, 2007). This mechanism seems particularly suited for the specific case of the quasi-rhythms of natural speech (Ding, Melloni, Zhang, Tian, & Poeppel, 2016), leading to the phenomenon referred to as brain-entrainment to speech: the ongoing oscillatory rhythmic neural activity becomes synchronous, or entrains, to the slow temporal fluctuation (envelope) of acoustic verbal stimuli (Ghitza & Greenberg, 2009; Giraud & Poeppel, 2012; Luo & Poeppel, 2007). Brain-entrainment to speech already starts at the speaker's side (Magrassi, Aromataris, Cabrini, Annovazzi-Lodi, & Moro, 2015; Pérez et al., 2017) due to the monitoring of one's own speech and/or due to the generation of the internal representations (motor and linguistics) preceding verbal output (Reznik, Henkin, Schadel, & Mukamel, 2014). Then, it takes place in the listener, linked to comprehension (Pelle, Gross, & Davis, 2013). Thus, at the basic physical level, B2B could be conceptualized as a situation in which one autonomous oscillator (speaker's brain) initiates an interaction (via speech) with another autonomous oscillator (listener's brain) leading to a mutual synchronization.

Albeit fostered and prompted by brain-entrainment to speech of the speaker and the listener, B2B not only occurs because of the joint entrainment to the speech envelope (Dai et al., 2018). The act of actively sharing relevant content through speech implies a form or partially disembodied interaction that goes beyond the physical verbal utterance (Pérez et al., 2017). During purposeful verbal information exchange, people intend to convey and receive meaning, and in order to facilitate

understanding and prediction of forthcoming speech, a form of mutual ‘alignment’ in attention to speech content is established between interlocutors (Friston & Frith, 2015a). The joint temporal fluctuations of attention resources associated with the timing and dynamics of the speech stream could be considered a type of joint attention necessary for successful communication and also an emergent property of the interactive scenarios contributing to the B2B phenomenon (Hari, Henriksson, Malinen, & Parkkonen, 2015). In general, it could be stated that shared attention to the act of communication and brain-entrainment to speech are two entangled factors (Alexandrou, Saarinen, Makela, Kujala, & Salmelin, 2017; Zion Golumbic et al., 2013) playing a critical role in B2B. Here, the question under scrutiny is whether modulating linguistic contextual factors in a ‘natural’ manner influences the patterns of the interbrain coupling.

Some linguistic contexts provide an ecologically valid scenario modulating the two mentioned factors. For instance, foreign language contexts represent a natural circumstance where intelligibility is hampered and the cognitive cost of producing and comprehending the language is higher than in native language contexts (Pérez, Carreiras, Gillon-Dowens, & Duñabeitia, 2015), thus requiring enhanced attention. On one hand, speech-entrainment is a causal mechanism that modulates intelligibility (Riecke, Formisano, Sorger, Baskent, & Gaudrain, 2018; Zoefel, Archer-Boyd, & Davis, 2018). In turn, intelligibility enhances entrainment (Buiatti, Pena, & Dehaene-Lambertz, 2009; Ding et al., 2016). From this positive correlation between brain-entrainment to speech and speech intelligibility it could be derived that the reduced intelligibility inherent to a foreign language (Munro & Derwing, 1995) would be associated with a reduced entrainment. In other words, since brain-entrainment to speech and speech intelligibility interact (Riecke et al., 2018; Zoefel et al., 2018), a decrease in the neural speech tracking would be expected for foreign language perception, at least for the low frequencies corresponding with parsing words from sentences. On the other hand, more cognitive resources should be recruited during foreign language processing to achieve a comprehension level comparable to that of a native language (see for review: Leow, 1997). In the specific case of speech perception, increased cognitive resources should be devoted in order to enhance speech tracking/comprehension to similar levels as in a native language. In fact, brain-entrainment to (natural) speech is enhanced by attention (Rimmele, Zion Golumbic, Schroger, & Poeppel, 2015; Zion Golumbic et al., 2013). In the case of speech production, it is known that the role of attention in monitoring native and nonnative languages is markedly different (Kormos, 2008). Hence, additional attentional resources should be allocated during foreign language production in order to generate the otherwise more difficult utterances (Rommers, Meyer, & Praamstra, 2017). Thus, more attentional resources should be deployed during nonnative speech perception and production. Altogether, it could be argued that using a foreign language reduces neural coupling to the speech and increases the cognitive load (attention) as compared to a native language (Perani & Abutalebi, 2005). Based on these assumptions, here we hypothesized that the B2B pattern would be different for linguistic interactions carried out using a native and a foreign language.

We used a two-person-neuroscience approach (Hari & Kujala, 2009; Schilbach et al., 2013) with electroencephalographic (EEG) hyperscanning (Babiloni & Astolfi, 2014) to simultaneously record the neural responses of two concomitant participants (a dyad) during a turn-taking verbal interaction in a native and foreign language. This experimental set-up, partially resembles a radio conversation in which

the roles of the speaker and the listener do not overlap while exchanging oral narratives and has been proven to be effective for the study of B2B (Pérez et al., 2017). In the current study, we employed this set-up to investigate the native-foreign contrast using dyads of native Spanish speakers who also had a good command of English. EEG phase synchronizations between brains were measured using the circular correlation coefficient (CCorr) (Jammalamadaka & Sengupta, 2001). A comparison of synchrony estimations between the brains, obtained for theta (3-7 Hz), alpha (8-12 Hz), and beta (13-30 Hz) bands, evaluated differences between the conditions. Those frequency bands contain the range of frequencies that are responsible for the cortical tracking at the timescale of phonemes, syllables, words and phrases (Keitel, Gross, & Kayser, 2018), and also have a functional role in attention deployment (Frey, Ruhnau, & Weisz, 2015). We expected interbrain synchronization differences across dyads associated with the language used. This result would suggest that a language choice for a verbal interaction affects the similarity in the neural patterns of the interlocutors, indicating the need for the inclusion of the linguistic factors in the theoretical accounts of interpersonal communication.

2. Methods

2.1. Participants

We report EEG data from 60 participants (40 females; mean age: 22.5, SD: 2.7 years, range: 18-30). They were recruited from the BCBL participant pool and received monetary compensation for their collaboration. All participants were right-handed as assessed by an adapted version of the Edinburgh Handedness Inventory (Oldfield, 1971), and self-reported adequate hearing, normal or corrected to normal vision, no neurological/psychiatric disorders, and no current drug use. All had completed high school, and most were undergraduates from the University of the Basque Country. None of them had any speech or reading disorder. Individual written informed consent was obtained before the experimental session. The study was conducted in accordance with the Declaration of Helsinki (1964) and approved by the BCBL Ethics Committee.

The linguistic competence in Spanish and English languages was estimated by an interview and a lexical decision test (LexTale) (de Bruin, Carreiras, & Duñabeitia, 2017). In the personal interview, all participants were classified as native Spanish speakers and received a score equal or above 3 (out of 5) in the part conducted in English, indicating at least basic fluency. The percentage of correct responses in the LexTale was markedly higher for Spanish (mean: 93.29, SD: 7.75) than for English (mean: 68.90, SD: 8.41), as expected given the difference in proficiency between the two tested. In addition, their first exposure to English occurred in all cases after the age of three.

Participants performed the study arranged in pairs (i.e. 30 dyads). Each dyad was composed of participants of the same gender (Cheng, Li, & Hu, 2015), equivalent age (with no more than five years of age difference), and they did not know each other before the experiment. We had three a-priori criteria to exclude participants' data: (i) technical issues during EEG acquisition such as saturation of the amplifiers, high electrode impedance or electrode detachment, (ii) participant discomfort with the experimental set-up, expressed overtly by them or perceived by the experimenter, and (iii) more than 25% of the EEG signal at any trial being irrecoverable due to large artifacts (e.g. laughing, sneezing,

coughing, yawning or body movements). Data from three additional dyads were excluded due to these criteria. In the case of the analysis involving the audio signal, data from four of the 30 dyads were not included due to technical issues with the audio recordings (i.e., N=26 dyads).

2.2. *Experimental Task*

The dyads of participants were required to alternate between the roles of speaker and listener while interchanging verbal information. Specifically, they were asked to talk and listen about preferences and opinions for six different common topics while following a short plot in the form of five questions that was presented on the screen. Figure 1 depicts an example of this core structure of the experimental task. First, one topic accompanied by a series of five questions was presented on the screen. A cue indicated which of the two participants in dyad had to speak and which had to listen. Thus, depending on the assigned role, participants had to either speak as freely and naturally as possible about the topic (Speaker role), or to listen while paying careful attention to the other's speech (Listener role). In each turn, the speaker had 120 seconds to talk. Information about the remaining time was presented on the screen to facilitate the structuring of the discourse. The countdown clock showed the remaining time at irregular intervals. After each trial, the roles were reversed. Thus, both participants performed the role of speaker and listener for each corresponding topic. In each topic/block, we also recorded a baseline trial of the same length in which both participants had to remain silent. The key manipulation in the experimental design was that half of the topics should be discussed in English (foreign language; English condition) whereas the other topics should be discussed in Spanish (native language; Spanish condition). The language to speak was cued by the language used in the text presented on the screen. Sports, travel and food were used as topics for the English conditions and movies, animals and music, for the Spanish condition. The presentation order of the conditions and topics were random, and so was the order for the roles of listener, speaker and silence in each block. Hence, the experimental task comprised six listening trials, six speaking trials and six silent trials for each participant of the dyad. Half of these trials belonged to the English condition and the other half to the Spanish condition. After completing the experiment, participants were asked to complete a paper and pencil evaluation concerning the information provided by their partner. It consisted of a 30-item recall test created following the specific questions (and language) that were displayed to guide the plot of the conversation for each of the topic, and consequently these were the same for all participants (e.g., "What's his/her favorite sport?"). Each participant scored their partner's responses to assess whether they had paid attention to their speech.

2.3. *Experimental setting*

The experimental procedure started with the experimenters introducing the pair of participants to each other. The experimenters and participants continued engaging in small talk while setting up the EEG cap and electrodes. Detailed instructions and explanations about the experiment (including an example) were provided. Participants were specifically asked to (i) be relaxed and still, avoiding body movements (e.g. gesticulation, nodding); (ii) speak continuously and without much pausing, vocalize clearly and avoid whispering and the use of interjections when cued to speak; and (iii) pay careful attention to the content of the partner's verbal production. They were also informed that a short evaluation about the content of the partner's verbal productions (post-scan questionnaire) would take place at the end of the

experiment and that they would receive a small monetary reward (€5) if responding correctly on at least 24 of the 30 questions.

Participants were seated side by side in a soundproof cabin, with a board placed between them to prevent them from seeing each other. A computer screen facing participants was placed at a distance of approximately one and a half meters, centred, and it was equally visible to the two participants. One microphone was used to record the verbal productions of both participants. The experiment was self-paced, thus, each participant was provided with a response pad that was used to trigger the start of each trial. Presentation on the screen of all instructions and stimuli was controlled by a custom-written program created and compiled with Experiment Builder® software (SR-Research, Ontario, Canada) that was run on a PC. The program also recorded the audio files and sent the triggers to the PC used for the EEG recordings.

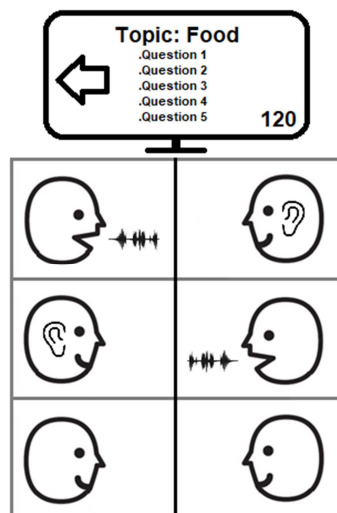


Figure 1. Schematic depiction of the experimental tasks. On each trial, the topic and a list of five questions to discuss, and an indication of whom should speak appeared on the screen. In the example, the topic is Food and the arrow points to the left, so the English language should be used and the left participant should be speaking. In the bottom-right corner, the time remaining in the trial was displayed at irregular intervals. Later on, the roles are reversed.

2.4. EEG hyperscanning recordings

Electrophysiological signals of each subject were acquired from a 32-channels BrainAmp Standard amplifier (Brain Products GmbH) with individual Reference and Ground electrodes. The two BrainAmps were connected to a USB 2 Adapter (BUA), and joint signals were monitored/recorded using one custom workspace in BrainVision Recorder. Elastic caps (EasyCap) mounted with 27 scalp Ag/AgCl electrodes were used to record the signal. Electrodes were placed according to the International 10-20 system and included Fp1/Fp2, F3/F4, F7/F8, FC1/FC2, FC5/FC6, C3/C4, T7/T8, CP1/CP2, CP5/CP6, P3/P4, P7/P8, O1/O2, Fz, Cz, and Pz. Ground and online reference electrodes were placed at AFz and FCz, respectively. Additional electrooculography (EOG) electrodes were placed at the external ocular canti (to monitor horizontal eye movements), and three electromyography (EMG) electrodes were placed at the right

cheek, as well as superior and inferior left orbicularis oris muscle (halfway between the centre and the corner of the mouth). Impedance measurements were checked by two individual 32-channels workspaces corresponding to each participant/amplifier. Inter-electrode impedances were set below 5 k Ω at the beginning of the experiment. Data were acquired at a sampling rate of 250 Hz. EEG markers were time-locked to the beginning of each trial.

2.5. EEG Preprocessing

The data were analyzed using the EEGLAB toolbox (Delorme & Makeig, 2004) v14.1.1 and custom programs, all running in MATLAB (version 2017a, The MathWorks Inc.). The recorded signal was high-pass filtered at 2 Hz and separated in the two recordings corresponding to each participant. Then, an adaptive mixture independent component analysis (AMICA) technique (Palmer, Kreutz-Delgado, Rao, & Makeig, 2007) was applied to the data from each participant. AMICA has been shown to maximize mutual information reduction and the dipolarity of scalp projections following decomposition (Delorme, Palmer, Onton, Oostenveld, & Makeig, 2012). The number of time points used to estimate the weighting matrix exceeded the minimum recommended amount for satisfactory decomposition in all dyads. Next, an equivalent dipole current source was fit to each IC using the DIPFIT toolbox of EEGLAB (Oostenveld & Oostendorp, 2002). Then, individual components accounting for blinks and saccades, heartbeat, muscle artifact or line noise were removed from the data. Component rejection was performed manually, guided by the following criteria: (i) the component's topography, (ii) the component's time-series, (iii) the component's power spectrum properties and (iv) properties of the dipole associated with each component: localization out of the head or close to eyes balls that was also associated with low variance (up to 15%). On average, we removed 4 components (range: 3–7, SD: 1.02). Next, an artifact subspace reconstruction (ASR) algorithm (T. Mullen et al., 2013; T. R. Mullen et al., 2015) adapted for EEGLAB software (clean_rawdata plugin) was implemented to remove high amplitude artifact from the EEG. ASR transforms a sliding window of EEG data with principal component analysis to identify channels and times of high variance by statistical comparison with clean EEG data containing minimal artifact. The clean data used as calibration data for ASR were automatically found inside each EEG recording. Channels that show variance above a threshold compared to calibration data were eliminated. Corrupted subspaces of multiple channels were reconstructed from neighbouring channels using a mixing matrix that is computed from the covariance matrix of the calibration data. In this study, a sliding window of 500 ms and a variance threshold of 3 SD were used. Those portions of the data that were not possible to reconstruct due to the presence of multiple artifacts were marked to be removed in a posterior step (in seconds: $M_{\text{English}} = 1$, $SD_{\text{English}} = 1.4$; $M_{\text{Spanish}} = 1.7$, $SD_{\text{Spanish}} = 2$; $t(58) = 1.5$, $p = .14$). Then, bad channels were interpolated. Subsequently, the continuous EEG signal was filter in the frequency band of interest, and 122 seconds epochs were extracted beginning one second prior to the onset of each stimulus. The instantaneous phase was estimated using the Hilbert transform, creating circular variables. Data from the first and last seconds and those time points not possible to reconstruct due to the presence of multiple artifacts were excluded. The CCorr was calculated over every possible combination of dyad's EEG electrodes (total of 729) in each trial by using the Circular Statistics Toolbox (Berens, 2009). Next, the calculated CCorr was normalized by Fisher's Z transformation and converted to absolute values. Normalized values obtained in the Silence condition were subtracted from those

obtained in each corresponding topic (Jiang et al., 2012; Pérez et al., 2017). If the coupling value for the Silence condition was larger than the obtained for the speaking and listening, subtraction was set to zero. This baseline normalization allows to reduce/eliminate non-specific kinds of engagement that may result from physical cohabitation as well as spurious synchronizations that may appear in hyperscanning EEG data (Burgess, 2013). Finally, the data were collapsed across topics for the Spanish and English conditions. This resulted in normalized estimations of synchronization in each frequency band, between the 27 channels of the Speaker and the 27 channels of the Listener. The Speaker-Listener situation is twofold for each member of the dyad, yielding independent observations of Speaker and Listener roles for all the participants.

2.6. Speech Processing

We computed the amplitude envelope of speech signals using a procedure similar to that described (Chandrasekaran, Trubanova, Stillitano, Caplier, & Ghazanfar, 2009) and employed in preceding studies (Gross et al., 2013; Park, Kayser, Thut, & Gross, 2016). Using the Chimera toolbox (Smith, Delgutte, & Oxenham, 2002), nine frequency bands in the range 100–10000 Hz that were equidistant on the cochlear map were constructed. Speech signals were band-pass filtered in these bands, and a Hilbert transform was applied to obtain amplitude envelopes for each of them. A wideband amplitude envelope was obtained by averaging envelopes across bands. The resulting envelopes were downsampled to 250 Hz for further analysis.

2.7. Statistical Analysis

A nonparametric bootstrap-based t-test method from EEGLAB's Resampling Statistical Toolkit was used to compare interbrain synchronization and brain to speech synchronization between language conditions. The bootstrap test is a distribution-free test and it does not require any assumptions about the correlation structure of the data. The number of random sampling was set to 10,000. An FDR correction (Benjamini & Yekutieli, 2001) for multiple comparisons ($q=0.05$) was applied to the p values obtained. Statistically significant effects are interpreted as showing differential B2B/speech tracking between native and foreign languages. Note that the experimental within-subjects design and corresponding statistics directly compared conditions and performed stringent corrections for multiple comparisons and spurious synchronizations.

3. Results

3.1. Brain-to-speech synchronization

First, the phase synchronization between the EEG signal and the speech audio envelope was compared between the Spanish and English conditions at the three frequency bands, separately, for the Listener and the Speaker. The heads in the left panel of Figure 2 show the topographical distribution of the t-values from nonparametric permutation tests. The table on the right shows, for both roles and frequency bands, the electrodes at which the difference was statistically significant ($p_{FDR}<.05$), by coloring in green a brain-to-speech coupling increased for the Foreign condition (English) and in red the

coupling increased for the Native condition (Spanish). The statistics of these comparisons are included in Supplementary Material 1 (all $df=55$).

Theta: Differential neural entrainment to the amplitude envelope was found for the Listener, reflected by increased coupling at Fz and FC6 for Spanish and English language, respectively.

Alpha: No statistically significant differences emerged.

Beta: Both the Speaker and the Listener showed increased brain-to-speech coupling at electrodes T8, P7 and FC5 for the Speaker and at P4 for the Listener. Beta band is supposed to reflect the close relationship between language comprehension and motor functions (Weiss & Mueller, 2012).

In general, these results suggest that neural tracking of a speech signal could variate depending on the language used, with a reduction in tracking in the less proficient language. This novel indication of differential brain-entrainment to speech for native and foreign languages already suggest that B2B patterns would be also differential. This is analysed in the next section.

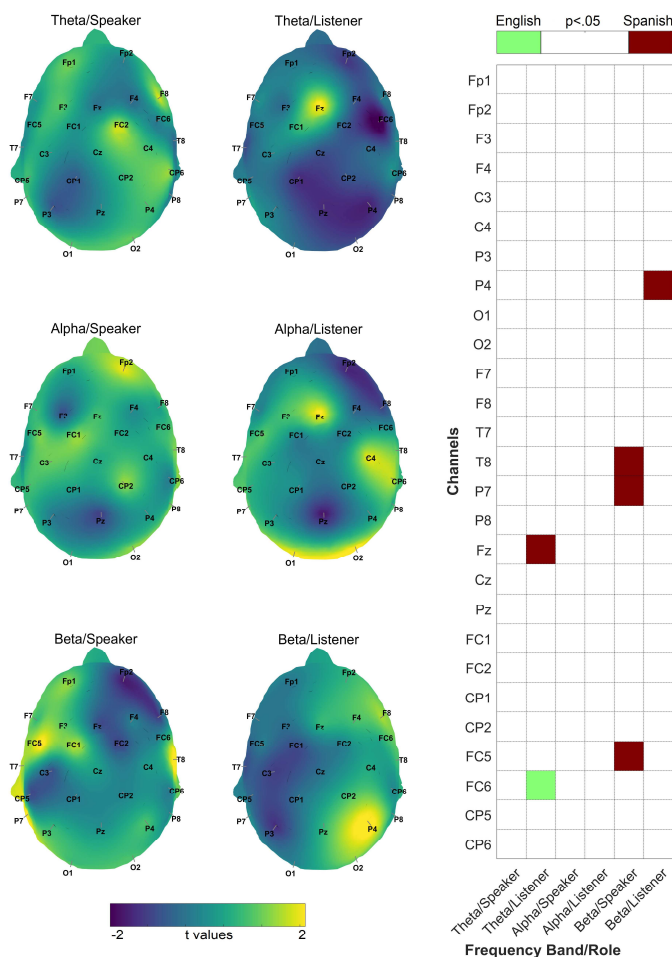


Figure 2. Group-level differences in brain entrainment to the speech envelope between the English and Spanish language conditions, for theta, alpha and beta bands, for the Speaker and the Listener's roles. The head plots show the topological distribution of the t-values from nonparametric permutation tests. The different colors represent the scale from -2 to 2. The matrix on the right shows the exact electrodes where the differences between condition are statistically different ($p_{FDR} < 0.05$) by coloring in green a brain-to-speech coupling increased for the Foreign condition (English) and in red the coupling increased for the Native condition (Spanish). Rows represent the electrodes and columns represent the role-frequency band. Significantly different brain-to-speech synchronization is mainly due to an increase in tracking for the Native condition.

3.2. Brain-to-brain synchronization

The baseline-corrected phase synchronization calculated for all Listener/Speaker electrode pair combinations of the EEG signals (i.e. 27×27 electrodes) was compared between the Spanish and English conditions for theta, alpha, and beta frequency bands for all the participants ($N=60$, all $df=59$). Figure 3 presents dual head plots and a heat map image of the differential patterns of interbrain phase synchronization between language conditions that were statistically significantly different ($p_{FDR} < .05$). Both representations depict the same data in complementary ways to facilitate the interpretation of the exact Listener/Speaker electrode pairs at which the differences were observed. The colours green and red are used to specify if the synchronization is larger for the native, Spanish condition (red) or the foreign, English condition (green). In the dual head plot, lines linking the electrodes of the Listener and the Speaker represent the differences. In the heat map, the intersection between the electrode pairs of the Listener (rows) and the Speaker (columns) are coloured to show the electrodes pairs at which there were significant differences between conditions. Supplementary Material 2 contains all exact p-values (uncorrected) and the corresponding t-values associated to these figures.

Theta and Beta: No statistically significant differences between conditions resulted after the FDR correction.

Alpha: Language condition differentially modulated the between-brain coupling at 14 electrodes-pairs, broadly distributed across the scalp of the Listener and the Speaker: 5 show a larger synchronization for the native language context (Spanish), and 9 for the foreign language context (English). Note that electrodes-pairs showing an effect do not include channels identified as bad in more than the half of the participants (namely, Fp1 and F7). The coupling strength between brain signals is increased in both the native language context and the foreign language context. That is, relative to English, we saw increased coupling at Spanish, but also the opposite case.

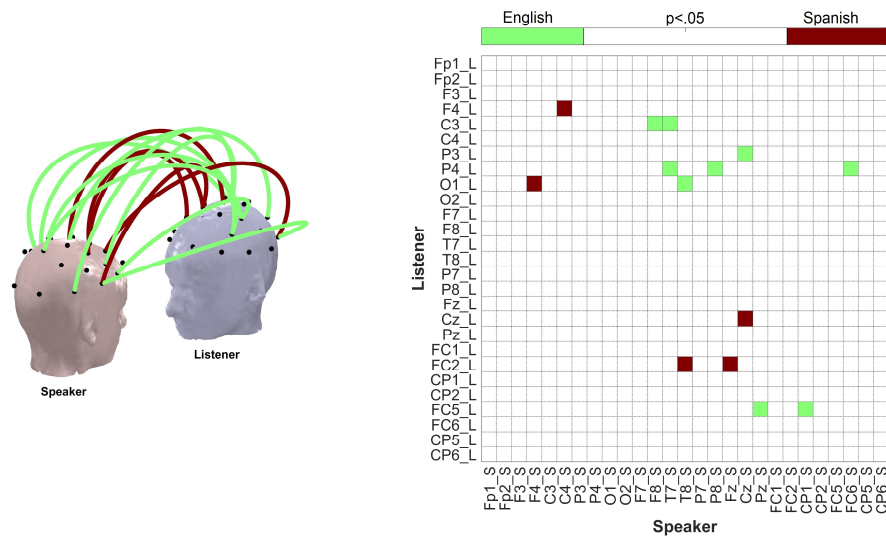


Figure 3. Group-level differences in interbrain synchronization patterns between the Foreign and Native language conditions for the alpha band. The dual head plot and the matrix constitute equivalent representations. The coloured lines are connecting those electrodes pairs from the Speaker (red head) and the Listener (blue head) showing statistically significant differences between conditions ($p_{FDR} < .05$). The same results are indicated in the matrix by colouring the intersection between Speaker's electrodes (columns) and Listener's electrodes (rows). Interbrain coupling increased for the Native condition (Spanish) is indicated in red and increased for the Foreign condition (English) is indicated in green. Significantly different interbrain synchronization is evident between the Native and Foreign language conditions in a broad scalp topography.

3.3. Behaviour

Overall performance in the 30-item recall test was markedly high (mean: 96.83%, SD: 5.22, range: 80-100). The high level of recall did not depend on the language in which the conversation about each topic was held since the percentages of correctly recalled items did not significantly vary across language blocks ($M_{\text{English}}: 97.33$, $SD_{\text{English}}: 5.65$; $M_{\text{Spanish}}: 96.44$, $SD_{\text{Spanish}}: 6.78$; $t(59)=1.03$, $p=.31$). The additional analysis confirmed that the performance in the different individual topics was similar, with average accuracy levels ranging from 93.33% to 98.67%. This indicate the topics elicited similar interest.

4. Discussion

Prior studies using the hyperscanning technique have already demonstrated differences in the coupling patterns between brains due to factors like gaze (Lachat, Hugueville, Lemarechal, Conty, & George, 2012; Leong et al., 2017), cooperation (Astolfi et al., 2010), movement (Dumas, Nadel, Soussignan, Martinerie, & Garnero, 2010), or intention (Tang et al., 2016). Here we demonstrate that differences in interbrain phase synchronization could be also due to the linguistic context where the exchange of verbal information is set, as illustrated by differences for native and foreign language contexts.

As it could be expected from the studies showing that neural coupling is significantly diminished for an unintelligible language (Stephens et al., 2010), there is a relative increase in interbrain coupling when the speaking and listening takes place in the more intelligible (native) language. However, a relative enhancement of the B2B coupling also takes place in a foreign language context. These results suggest that differences are not exclusively due to a decrease in the quality of the communication when the foreign language is used, given that under this assumption no signs for an enhanced interbrain coupling would have been expected for the foreign language as compared to the native language context. Moreover, considering the behavioral results showing that the item recall test was equivalent across the different language conditions, an explanation purely based on the intelligibility of foreign language speech cannot account for the general pattern of results here presented. Instead, this twofold directionality of the effects indicates that interbrain neural alignment depends on the language context and the demands of speaking and listening in this context.

The finding of dissimilar B2B patterns associated to different linguistic contexts could be interpreted according to the framework of the interactive linguistic alignment (ILA) theory (see Pickering & Garrod, 2004). According to this theory, during verbal communication, production and comprehension processes become aligned at different levels: phonetic, phonological, lexical, syntactic and semantic. In fact, the linguistic operations needed to understand and produce an utterance rely to a great extent on similar networks of the brain (Menenti, Gierhan, Segaert, & Hagoort, 2011). Thus, interpersonal linguistic alignment during speaking and listening will be supported by a similar underlying neural configuration (and timing) at each of those linguistic levels mentioned above. Importantly, the neural configuration of each specific linguistic feature will depend on the language system or code used. For illustration purposes, consider the phonological/articulatory alignment for the word 'snake' being produced and perceived in English. Even under the assumption of the existence of subject-invariant semantic representations (Mahon & Hickok, 2016), the phonological/articulatory alignment would be different for the corresponding Spanish translation equivalent ('serpiente'). In an extreme positioning of this view, it could be also considered that linguistic features may not be shared across languages at all, and that bilinguals exhibit language-dependent neuroanatomical and semantic representations (García-Pentón, Fernández García, Costello, Duñabeitia, & Carreiras, 2016). In other words, the functional organization (i.e., the neural activation patterns) related to the processing of distinct linguistic features may be different for different language systems (Pérez, Gillon-Dowens, et al., 2015), consequently impacting differently on the neural alignment between interlocutors across languages.

In the current study, differential B2B synchronization or alignment while speaking and listening in native and foreign languages was evident in the alpha frequency band. Increased interbrain coupling results from a direct statistical comparison between conditions and it is interpreted as an indication of a greater similarity between the neurophysiological activity of two different brains. This means that the current results speak for a more similar alpha oscillatory activity in one language vs. the other, but not necessarily for an increase in alpha activity. Alpha oscillations are the prevalent rhythm of the brain, being linked to main cognitive functions such as attention (Klimesch, 2012), working memory (Wilsch, Henry, Herrmann, Maess, & Obleser, 2015) or decision making (Cohen, Elger, & Fell, 2009). In fact, alpha oscillations seem to play a major role in functional inhibition (Jensen & Mazaheri, 2010), hindering the

processing capabilities of a given area of the brain. A corollary to that idea is that optimal task performance will correlate with alpha activity in task-irrelevant areas (Jensen & Mazaheri, 2010). Interestingly, the topography of the effects denoting increased alignment in alpha is more broadly distributed over the scalp and includes more pairs of electrodes in a foreign language contexts than in the native language, in line with the requirement of an increased auditory selective inhibition due to the challenging conversational situation represented foreign language exchanges of information (Strauß, Kotz, Scharinger, & Obleser, 2014; Strauß, Wostmann, & Obleser, 2014). Besides, alpha is one of the dominant rhythms in motor areas (Pineda, 2005), and the specific pattern of increased brain-to-brain entrainment in the native language is clustered in fronto-central areas. This could tentatively be taken as evidence supporting a larger similarity between speakers and listeners at the sensory-motor levels of speech production and perception when the interlocutors are expert users of a given language. This observation is supported by the described shared ventral premotor substrate for real-life speech production and perception (Glanz Ijina et al., 2018) and the central alpha activity influencing speech tracking in a native language (Keitel, Ince, Gross, & Kayser, 2017). Overall, the differential B2B alignment found in the alpha (and not theta) band is in line with the evidence supporting the role of alpha in linguistic processing (Obleser & Weisz, 2012).

We further interpret this differential B2B pattern in the alpha band as a result from the different network configurations operating because of the language-idiosyncratic joint attentional fluctuations. In our view, 'purely attentional' alpha oscillations covary for speech perception and speech production depending on the language context. Joint attention is a ubiquitous element of theories of successful interpersonal communication (Schirmer, Meck, & Penney, 2016). Successful communication requires coordinated attention between interlocutors. In the case of the verbal communication, this tight relationship in the attentional fluctuations is established due to interactive prediction processes between the interlocutors (Friston & Frith, 2015b). For different languages, there is a change in the probabilistic nature of linguistic behavior, the structure of thought processes, and the perceptual/motor biases (among others), leading interlocutors to prioritize different aspects of events in narrative discourse (Beckner et al., 2009). In the context of the current study, and considering that participants were not balanced bilinguals, it seems reasonable to assume fundamental differences between both language settings related to phonemic and prosodic awareness. This way, language proficiency could be automatically imposing particular attentional demands and strategies (i.e., attending smaller chunks of information) when speaking and listening in a foreign language, yielding a direct impact on attentional alignment (Pérez, Carreiras, et al., 2015). In conclusion, we claim that communication in a foreign language would need reshaping the attentional strategies to achieve a mutual understanding resulting in a distinctive fluctuation of joint attention reflected in the alpha band.

Although the focus of our study was on the interbrain synchronization between language contexts, we also noted difference in brain-to speech entrainment to native and foreign languages. Considering that all participants were heavily accented nonnative speakers of English as a foreign language, this could have modulated the brain-to-speech entrainment patterns, adding complexity to the interpretation. However, quite interestingly, no brain-to-speech effects emerged in alpha band, where the main brain-to-brain differences were found, suggesting that these two entrainment effects are somewhat

independent from each other. Although speech intelligibility and brain-entrainment to speech interact reciprocally depending on attention, the effects of brain-to-speech entrainment have been found to be independent of brain-to-brain synchronization. Furthermore, the specific timeline of speech tracking is different for speakers and listeners, highlighting also temporal differences between B2B and brain-to-speech entrainment depending on the conversational role being played (Stephens et al., 2010). Thus, while B2B and brain-to-speech entrainment relate to seemingly conceptually related elements, establishing a direct link between them is not supported by current evidence, suggesting that one does not emerge as a direct transitive property of the other.

It remains unclear whether factors other than linguistic alignment, joint attention and brain-entrainment to speech could yield or interact to lead to the differential interbrain synchronization patterns. One could also attribute the differential entrainment pattern to differences in the emotional processing associated with foreign language communication (Lev-Ari, Ho, & Keysar, 2018) or because using a foreign language reduces mental imagery (Hayakawa & Keysar, 2018). Emotions elicit synchronization in brain activity across individuals (Nummenmaa et al., 2014) and the degree of empathy correlates with the brain-to-brain coupling (Goldstein, Weissman-Fogel, Dumas, & Shamay-Tsoory, 2018). Less emotional engagement in foreign language contexts could be shaping the interbrain synchronization patterns during verbal interactions. Nonetheless, this is admittedly just one of many possible additional factors, and the additional underlying forces (see also Mu, Guo, & Han, 2016) driving the differential B2B pattern to the foreign language remain as open questions for future research.

Brain entrainment to speech is considered as a mechanistic component of speech encoding serving speech segmentation and parsing at different timescales (e.g., words and syllables). In the same way, brain-to-brain entrainment while speaking and listening could be an emergent mechanism enabling to convey information across brains by fostering mutual intelligibility, coordinated (joint) attention and the linguistic alignment needed for successful communication between individuals. Nonetheless, and as a cautionary note, we want to stress that the characterization of B2B as a mechanism that underlies human communication is still premature, given that its specific internal organization as well as its causal role need to be first identified (Craver & Tabery, 2017). Although previous results indicate that B2B contributes to successful communication (Stephens et al., 2010), this does not necessarily imply that interbrain synchronization is sufficient or even required for successful communication. B2B remains as a measurable reflection of the underlying neural computations that underpin shared cognitive processes (Dikker et al., 2017), and future studies on the interactions between B2B and comprehension using techniques that allow for direct stimulation and disruption of the processing in the critical areas (Zoefel & Davis, 2017) will clarify if neural alignment is indeed a mechanism underlying the phenomenon of verbal communication.

It is worth noting that any synchronization effects obtained in the current experiment have been solely driven by the linguistic interaction, and not by embodied interactions like lip movements or any kind non-verbal communication. In our view, avoiding visual contact between the interlocutors is beneficial for the purposes of experimental neurolinguistic studies aimed at tracking synchronous neural activity in response to verbal interactions, since visual cues (facial, gestural) not only supplement but sometimes even override the speech signal. Also, interlocutors speak more intelligibly when they cannot see one

another (Bruce, Braidwood, & Newton, 2013; Pate & Goldwater, 2015), and this could favor the purposes of studies such as the current one. Moreover, the turn-taking strategy used in our experimental setup resembles ‘formal’ real life situations, such as interviews and panel discussions, and other scenarios in which the timing of the interactions (and the interaction itself) is constrained by social rules and expectations. Thus, we believe that the paradigm used here could be considered both efficient and ecologically valid for the study of the role of linguistic interactions. Nonetheless, we stress that face-to-face conversations improve the quality of communication, and this could yield different B2B patterns. Future methodological advances will be of help to explore B2B entrainment in different communicative scenarios.

In sum, it seems clear that the type or nature of the linguistic context impacts on the alignment of neural activity between speakers and listeners, at least for the case tested here with languages in which participants had dissimilar proficiency. These interbrain synchronization effects highlight some mechanisms that occur in linguistic interactions that cannot be otherwise captured by assessing individual brains (Hasson, Ghazanfar, Galantucci, Garrod, & Keysers, 2012). Findings here open doors to the study of the role of linguistic factors as mediators of social interactions using a multi-person perspective (Evans, 2013). We propose that the study of neurolinguistics more centered on the neural changes resulting from the joint communicative intention will help uncovering the mechanisms and processes involved in interpersonal interactions, improving our understanding of language processing in more ecologically valid settings. As illustrated by the present work, it is only by taking into account the multiple dimensions of factors that influence two-way communication that we will be able to develop comprehensive theoretical accounts of how humans communicate with one another.

Code availability

The code that support the findings of this study is available upon request to A.P.

Data availability

The data that support the findings of this study is property of the BCBL institution. For data access request please contact: info@bcbl.eu.

Acknowledgment

We thank Rosa M. Fernández, Jo-Anne Blanchet and Gino Ruffo for their support at different stages of the study. We also thank Dr. Blair Armstrong for the revision of the manuscript. The authors acknowledge financial support from the Spanish Ministry of Economy and Competitiveness through the “Severo Ochoa Programme for Centres/Units of Excellence in R&D” (SEV-2015-490) and the PSI2015-65689-P grants.

Author Contributions

A.P., G.D., J.A.D. conceived and designed the experiment. M.K. collected and organized the data. A.P. processed the data. A.P., G.D. and J.A.D. analyzed the data and interpreted the results. A.P. drafted the

manuscript and the responses to the reviewers' comments, which were reviewed, adapted and approved by all authors.

References

- Ahn, S., Cho, H., Kwon, M., Kim, K., Kwon, H., Kim, B. S., . . . Jun, S. C. (2017). Interbrain phase synchronization during turn-taking verbal interaction-a hyperscanning study using simultaneous EEG/MEG. *Hum Brain Mapp*. doi:10.1002/hbm.23834
- Alexandrou, A. M., Saarinen, T., Makela, S., Kujala, J., & Salmelin, R. (2017). The right hemisphere is highlighted in connected natural speech production and perception. *Neuroimage*, 152, 628-638. doi:10.1016/j.neuroimage.2017.03.006
- Astolfi, L., Toppi, J., De Vico Fallani, F., Vecchiato, G., Salinari, S., Mattia, D., . . . Babiloni, F. (2010). Neuroelectrical hyperscanning measures simultaneous brain activity in humans. *Brain Topogr*, 23(3), 243-256. doi:10.1007/s10548-010-0147-9
- Babiloni, F., & Astolfi, L. (2014). Social neuroscience and hyperscanning techniques: past, present and future. *Neurosci Biobehav Rev*, 44, 76-93. doi:10.1016/j.neubiorev.2012.07.006
- Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., . . . Schoenemann, T. (2009). Language Is a Complex Adaptive System: Position Paper. *Lang Learn*, 59(s1), 1-26. doi:10.1111/j.1467-9922.2009.00533.x
- Benjamini, Y., & Yekutieli, D. (2001). The Control of the False Discovery Rate in Multiple Testing under Dependency. *The Annals of Statistics*, 29(4), 1165-1188.
- Berens, P. (2009). CircStat: A MATLAB Toolbox for Circular Statistics. *J Stat Softw*, 1(10). doi:10.18637/jss.v031.i10
- Bruce, C., Braidwood, U., & Newton, C. (2013). 'All the better for not seeing you': effects of communicative context on the speech of an individual with acquired communication difficulties. *J Commun Disord*, 46(5-6), 475-483. doi:10.1016/j.jcomdis.2013.08.002
- Buiatti, M., Pena, M., & Dehaene-Lambertz, G. (2009). Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses. *Neuroimage*, 44(2), 509-519. doi:10.1016/j.neuroimage.2008.09.015
- Burgess, A. P. (2013). On the interpretation of synchronization in EEG hyperscanning studies: a cautionary note. *Front Hum Neurosci*, 7, 881. doi:10.3389/fnhum.2013.00881
- Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., & Ghazanfar, A. A. (2009). The natural statistics of audiovisual speech. *PLoS Comput Biol*, 5(7), e1000436. doi:10.1371/journal.pcbi.1000436
- Cheng, X., Li, X., & Hu, Y. (2015). Synchronous brain activity during cooperative exchange depends on gender of partner: A fNIRS-based hyperscanning study. *Hum Brain Mapp*, 36(6), 2039-2048. doi:10.1002/hbm.22754
- Cohen, M. X., Elger, C. E., & Fell, J. (2009). Oscillatory activity and phase-amplitude coupling in the human medial frontal cortex during decision making. *J Cogn Neurosci*, 21(2), 390-402. doi:10.1162/jocn.2008.21020
- Craver, C., & Tabery, J. (2017). Mechanisms in Science. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2017 ed.). Stanford University: Metaphysics Research Lab.

- Dai, B., Chen, C., Long, Y., Zheng, L., Zhao, H., Bai, X., . . . Lu, C. (2018). Neural mechanisms for selectively tuning in to the target speaker in a naturalistic noisy situation. *Nat Commun*, 9(1), 2405. doi:10.1038/s41467-018-04819-z
- de Bruin, A., Carreiras, M., & Duñabeitia, J. A. (2017). The BEST Dataset of Language Proficiency. *Front Psychol*, 8, 522. doi:10.3389/fpsyg.2017.00522
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods*, 134(1), 9-21. doi:10.1016/j.jneumeth.2003.10.009
- Delorme, A., Palmer, J., Onton, J., Oostenveld, R., & Makeig, S. (2012). Independent EEG sources are dipolar. *PLoS One*, 7(2), e30135. doi:10.1371/journal.pone.0030135
- Dikker, S., Silbert, L. J., Hasson, U., & Zevin, J. D. (2014). On the same wavelength: predictable language enhances speaker-listener brain-to-brain synchrony in posterior superior temporal gyrus. *J Neurosci*, 34(18), 6267-6272. doi:10.1523/jneurosci.3796-13.2014
- Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., . . . Poeppel, D. (2017). Brain-to-Brain Synchrony Tracks Real-World Dynamic Group Interactions in the Classroom. *Curr Biol*, 27(9), 1375-1380. doi:10.1016/j.cub.2017.04.002
- Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nat Neurosci*, 19(1), 158-164. doi:10.1038/nn.4186
- Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., & Garnero, L. (2010). Inter-brain synchronization during social interaction. *PLoS One*, 5(8), e12166. doi:10.1371/journal.pone.0012166
- Evans, N. (2013). On projecting grammatical persons into social neurocognition: a view from linguistics. *Behav Brain Sci*, 36(4), 419-420. doi:10.1017/s0140525x12001896
- Frey, J. N., Ruhnu, P., & Weisz, N. (2015). Not so different after all: The same oscillatory processes support different types of attention. *Brain Res*, 1626, 183-197. doi:10.1016/j.brainres.2015.02.017
- Friston, K., & Frith, C. (2015a). A Duet for one. *Conscious Cogn*, 36, 390-405. doi:10.1016/j.concog.2014.12.003
- Friston, K., & Frith, C. (2015b). Active inference, communication and hermeneutics. *Cortex*, 68, 129-143. doi:10.1016/j.cortex.2015.03.025
- García-Pentón, L., Fernández García, Y., Costello, B., Duñabeitia, J. A., & Carreiras, M. (2016). The neuroanatomy of bilingualism: how to turn a hazy view into the full picture. *Lang Cogn Neurosci*, 31(3), 303-327. doi:10.1080/23273798.2015.1068944
- Ghitza, O., & Greenberg, S. (2009). On the possible role of brain rhythms in speech perception: intelligibility of time-compressed speech with periodic and aperiodic insertions of silence. *Phonetica*, 66(1-2), 113-126. doi:10.1159/000208934
- Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nat Neurosci*, 15(4), 511-517. doi:10.1038/nn.3063
- Glanz Iljina, O., Derix, J., Kaur, R., Schulze-Bonhage, A., Auer, P., Aertsen, A., & Ball, T. (2018). Real-life speech production and perception have a shared premotor-cortical substrate. *Sci Rep*, 8(1), 8898. doi:10.1038/s41598-018-26801-x
- Goldstein, P., Weissman-Fogel, I., Dumas, G., & Shamay-Tsoory, S. G. (2018). Brain-to-brain coupling during handholding is associated with pain reduction. *Proc Natl Acad Sci U S A*, 115(11), E2528-e2537. doi:10.1073/pnas.1703643115
- Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., & Garrod, S. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biol*, 11(12), e1001752. doi:10.1371/journal.pbio.1001752
- Hari, R., Henriksson, L., Malinen, S., & Parkkonen, L. (2015). Centrality of Social Interaction in Human Brain Function. *Neuron*, 88(1), 181-193. doi:10.1016/j.neuron.2015.09.022

- Hari, R., & Kujala, M. V. (2009). Brain basis of human social interaction: from concepts to brain imaging. *Physiol Rev*, 89(2), 453-479. doi:10.1152/physrev.00041.2007
- Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends Cogn Sci*, 16(2), 114-121. doi:10.1016/j.tics.2011.12.007
- Hayakawa, S., & Keysar, B. (2018). Using a foreign language reduces mental imagery. *Cognition*, 173, 8-15. doi:10.1016/j.cognition.2017.12.010
- Jammalamadaka, S. R., & Sengupta, A. (2001). *Topics in circular statistics*. Singapore: World Scientific.
- Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Front Hum Neurosci*, 4, 186. doi:10.3389/fnhum.2010.00186
- Jiang, J., Dai, B., Peng, D., Zhu, C., Liu, L., & Lu, C. (2012). Neural synchronization during face-to-face communication. *J Neurosci*, 32(45), 16064-16069. doi:10.1523/jneurosci.2926-12.2012
- Kawasaki, M., Yamada, Y., Ushiku, Y., Miyauchi, E., & Yamaguchi, Y. (2013). Inter-brain synchronization during coordination of speech rhythm in human-to-human social interaction. *Sci Rep*, 3, 1692. doi:10.1038/srep01692
- Keitel, A., Gross, J., & Kayser, C. (2018). Perceptually relevant speech tracking in auditory and motor cortex reflects distinct linguistic features. *PLoS Biol*, 16(3), e2004473. doi:10.1371/journal.pbio.2004473
- Keitel, A., Ince, R. A. A., Gross, J., & Kayser, C. (2017). Auditory cortical delta-entrainment interacts with oscillatory power in multiple fronto-parietal networks. *Neuroimage*, 147, 32-42. doi:10.1016/j.neuroimage.2016.11.062
- Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends Cogn Sci*, 16(12), 606-617. doi:10.1016/j.tics.2012.10.007
- Kormos, J. (2008). The Role of Attention in Monitoring Second Language Speech Production. *Lang Learn*, 50(2), 343-384. doi:10.1111/0023-8333.00120
- Kuhlen, A. K., Allefeld, C., & Haynes, J. D. (2012). Content-specific coordination of listeners' to speakers' EEG during communication. *Front Hum Neurosci*, 6, 266. doi:10.3389/fnhum.2012.00266
- Lachat, F., Hugueville, L., Lemarechal, J. D., Conty, L., & George, N. (2012). Oscillatory Brain Correlates of Live Joint Attention: A Dual-EEG Study. *Front Hum Neurosci*, 6, 156. doi:10.3389/fnhum.2012.00156
- Lakatos, P., Chen, C. M., O'Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron*, 53(2), 279-292. doi:10.1016/j.neuron.2006.12.011
- Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*, 320(5872), 110-113. doi:10.1126/science.1154735
- Leong, V., Byrne, E., Clackson, K., Georgieva, S., Lam, S., & Wass, S. (2017). Speaker gaze increases information coupling between infant and adult brains. *Proc Natl Acad Sci USA*, 114(50), 13290-13295. doi:10.1073/pnas.1702493114
- Leow, R. P. (1997). Attention, Awareness, and Foreign Language Behavior. *Lang Learn*, 47(3), 467-505. doi:10.1111/0023-8333.00017
- Lev-Ari, S., Ho, E., & Keysar, B. (2018). The Unforeseen Consequences of Interacting With Non-Native Speakers. *Top Cogn Sci*, 1-15. doi:10.1111/tops.12325
- Liu, Y., Piazza, E. A., Simony, E., Shewokis, P. A., Onaral, B., Hasson, U., & Ayaz, H. (2017). Measuring speaker-listener neural coupling with functional near infrared spectroscopy. *Sci Rep*, 7, 43293. doi:10.1038/srep43293
- Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, 54(6), 1001-1010. doi:10.1016/j.neuron.2007.06.004

- Magrassi, L., Aromataris, G., Cabrini, A., Annovazzi-Lodi, V., & Moro, A. (2015). Sound representation in higher language areas during language generation. *Proc Natl Acad Sci U S A*, 112(6), 1868-1873. doi:10.1073/pnas.1418162112
- Mahon, B. Z., & Hickok, G. (2016). Arguments about the nature of concepts: Symbols, embodiment, and beyond. *Psychon Bull Rev*, 23(4), 941-958. doi:10.3758/s13423-016-1045-2
- Menenti, L., Gierhan, S. M., Segaert, K., & Hagoort, P. (2011). Shared language: overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. *Psychol Sci*, 22(9), 1173-1182. doi:10.1177/0956797611418347
- Mu, Y., Guo, C., & Han, S. (2016). Oxytocin enhances inter-brain synchrony during social coordination in male adults. *Soc Cogn Affect Neurosci*, 11(12), 1882-1893. doi:10.1093/scan/nsw106
- Mullen, T., Kothe, C., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., . . . Jung, T. P. (2013). Real-time modeling and 3D visualization of source dynamics and connectivity using wearable EEG. *Conf Proc IEEE Eng Med Biol Soc*, 2013, 2184-2187. doi:10.1109/embc.2013.6609968
- Mullen, T. R., Kothe, C. A., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., . . . Cauwenberghs, G. (2015). Real-Time Neuroimaging and Cognitive Monitoring Using Wearable Dry EEG. *IEEE Trans Biomed Eng*, 62(11), 2553-2567. doi:10.1109/tbme.2015.2481482
- Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Lang Speech*, 38 (Pt 3), 289-306. doi:10.1177/002383099503800305
- Nozawa, T., Sasaki, Y., Sakaki, K., Yokoyama, R., & Kawashima, R. (2016). Interpersonal frontopolar neural synchronization in group communication: An exploration toward fNIRS hyperscanning of natural interactions. *Neuroimage*, 133, 484-497. doi:10.1016/j.neuroimage.2016.03.059
- Nummenmaa, L., Saarimäki, H., Glerean, E., Gotsopoulos, A., Jaaskelainen, I. P., Hari, R., & Sams, M. (2014). Emotional speech synchronizes brains across listeners and engages large-scale dynamic brain networks. *Neuroimage*, 102 Pt 2, 498-509. doi:10.1016/j.neuroimage.2014.07.063
- Obleser, J., & Weisz, N. (2012). Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cereb Cortex*, 22(11), 2466-2477. doi:10.1093/cercor/bhr325
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9(1), 97-113. doi:10.1016/0028-3932(71)90067-4
- Oostenveld, R., & Oostendorp, T. F. (2002). Validating the boundary element method for forward and inverse EEG computations in the presence of a hole in the skull. *Hum Brain Mapp*, 17(3), 179-192. doi:10.1002/hbm.10061
- Palmer, J. A., Kreutz-Delgado, K., Rao, B. D., & Makeig, S. (2007). *Modeling and Estimation of Dependent Subspaces with Non-radially Symmetric and Skewed Densities*. Paper presented at the 7th International Conference, ICA 2007, London, UK.
- Park, H., Kayser, C., Thut, G., & Gross, J. (2016). Lip movements entrain the observers' low-frequency brain oscillations to facilitate speech intelligibility. *eLife*, 5, e14521. doi:10.7554/eLife.14521
- Pate, J. K., & Goldwater, S. (2015). Talkers account for listener and channel characteristics to communicate efficiently. *J Mem Lang*, 78, 1-17. doi:10.1016/j.jml.2014.10.003
- Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cereb Cortex*, 23(6), 1378-1387. doi:10.1093/cercor/bhs118
- Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Curr Opin Neurobiol*, 15(2), 202-206. doi:10.1016/j.conb.2005.03.007
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behav Brain Sci*, 27(2), 169-190. doi:10.1017/S0140525X04000056
- Pineda, J. A. (2005). The functional significance of mu rhythms: translating "seeing" and "hearing" into "doing". *Brain Res Brain Res Rev*, 50(1), 57-68. doi:10.1016/j.brainresrev.2005.04.005

- Pérez, A., Carreiras, M., & Duñabeitia, J. A. (2017). Brain-to-brain entrainment: EEG interbrain synchronization while speaking and listening. *Sci Rep*, 7(4190). doi:10.1038/s41598-017-04464-4
- Pérez, A., Carreiras, M., Gillon-Dowens, M., & Duñabeitia, J. A. (2015). Differential oscillatory encoding of foreign speech. *Brain Lang*, 147, 51-57. doi:10.1016/j.bandl.2015.05.008
- Pérez, A., Gillon-Dowens, M., Molinaro, N., Iturria-Medina, Y., Barraza, P., García-Pentón, L., & Carreiras, M. (2015). Complex brain network properties in late L2 learners and native speakers. *Neuropsychologia*, 68, 209-217. doi:10.1016/j.neuropsychologia.2015.01.021
- Reznik, D., Henkin, Y., Schadel, N., & Mukamel, R. (2014). Lateralized enhancement of auditory cortex activity and increased sensitivity to self-generated sounds. *Nat Commun*, 5, 4059. doi:10.1038/ncomms5059
- Riecke, L., Formisano, E., Sorger, B., Baskent, D., & Gaudrain, E. (2018). Neural Entrainment to Speech Modulates Speech Intelligibility. *Curr Biol*, 28(2), 161-169.e5. doi:10.1016/j.cub.2017.11.033
- Rimmele, J. M., Zion Golumbic, E., Schroger, E., & Poeppel, D. (2015). The effects of selective attention and speech acoustics on neural speech-tracking in a multi-talker scene. *Cortex*, 68, 144-154. doi:10.1016/j.cortex.2014.12.014
- Rommers, J., Meyer, A. S., & Praamstra, P. (2017). Lateralized electrical brain activity reveals covert attention allocation during speaking. *Neuropsychologia*, 95, 101-110. doi:10.1016/j.neuropsychologia.2016.12.013
- Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behav Brain Sci*, 36(4), 393-414. doi:10.1017/s0140525x12000660
- Schirmer, A., Meck, W. H., & Penney, T. B. (2016). The Socio-Temporal Brain: Connecting People in Time. *Trends Cogn Sci*, 20(10), 760-772. doi:10.1016/j.tics.2016.08.002
- Schoot, L., Hagoort, P., & Segaert, K. (2016). What can we learn from a two-brain approach to verbal interaction? *Neurosci Biobehav Rev*, 68, 454-459. doi:10.1016/j.neubiorev.2016.06.009
- Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, 416(6876), 87-90. doi:10.1038/416087a
- Spiegelhalter, K., Ohlendorf, S., Regen, W., Feige, B., Tebartz van Elst, L., Weiller, C., . . . Tuscher, O. (2014). Interindividual synchronization of brain activity during live verbal communication. *Behav Brain Res*, 258, 75-79. doi:10.1016/j.bbr.2013.10.015
- Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. *Proc Natl Acad Sci U S A*, 107(32), 14425-14430. doi:10.1073/pnas.1008662107
- Strauß, A., Kotz, S. A., Scharinger, M., & Obleser, J. (2014). Alpha and theta brain oscillations index dissociable processes in spoken word recognition. *Neuroimage*, 97, 387-395. doi:10.1016/j.neuroimage.2014.04.005
- Strauß, A., Wostmann, M., & Obleser, J. (2014). Cortical alpha oscillations as a tool for auditory selective inhibition. *Front Hum Neurosci*, 8, 350. doi:10.3389/fnhum.2014.00350
- Tadic, B., Andjelkovic, M., Boshkoska, B. M., & Levnjajic, Z. (2016). Algebraic Topology of Multi-Brain Connectivity Networks Reveals Dissimilarity in Functional Patterns during Spoken Communications. *PLoS One*, 11(11), e0166787. doi:10.1371/journal.pone.0166787
- Tang, H., Mai, X., Wang, S., Zhu, C., Krueger, F., & Liu, C. (2016). Interpersonal brain synchronization in the right temporo-parietal junction during face-to-face economic exchange. *Soc Cogn Affect Neurosci*, 11(1), 23-32. doi:10.1093/scan/nsv092
- Weiss, S., & Mueller, H. M. (2012). "Too Many betas do not Spoil the Broth": The Role of Beta Brain Oscillations in Language Processing. *Front Psychol*, 3, 201. doi:10.3389/fpsyg.2012.00201
- Wilsch, A., Henry, M. J., Herrmann, B., Maess, B., & Obleser, J. (2015). Alpha Oscillatory Dynamics Index Temporal Expectation Benefits in Working Memory. *Cereb Cortex*, 25(7), 1938-1946. doi:10.1093/cercor/bhu004

- Zion Golumbic, E. M., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., . . . Schroeder, C. E. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". *Neuron*, 77(5), 980-991. doi:10.1016/j.neuron.2012.12.037
- Zoefel, B., Archer-Boyd, A., & Davis, M. H. (2018). Phase Entrainment of Brain Oscillations Causally Modulates Neural Responses to Intelligible Speech. *Curr Biol*, 28(3), 401-408.e5. doi:10.1016/j.cub.2017.11.071
- Zoefel, B., & Davis, M. H. (2017). Transcranial electric stimulation for the investigation of speech perception and comprehension. *Lang Cogn Neurosci*, 32(7), 910-923. doi:10.1080/23273798.2016.1247970

Supplementary Material 1. Statistics from the brain-to-speech entrainment. Document containing all electrodes exact p-values (uncorrected) and the corresponding t-values obtained from the comparison between the Spanish and English conditions at the three frequency bands, for the Listener and the Speaker.

Supplementary Material 2. Statistics from brain-to-brain entrainment. Document containing all electrodes-pairs' exact p-values (uncorrected) and the corresponding t-values obtained from the comparison between the Foreign and Native language conditions for the alpha band.