EEG Signal Analysis for the Detection of Spoken Language Comprehension

Sergio Guillén(1), Lorenzo J. Tardón(1), Ana M. Barbancho(1), Irene De-Torres(2), Isabel Barbancho(1)
sergiogj@ic.uma.es, ltg@uma.es, abp@uma.es, detorres.irene@gmail.com, ibp@uma.es
(1)ATIC Research Group, E.T.S.I. Telecomunicación, Universidad de Málaga, E-29071, Málaga, Spain.
(2)Hospital Regional Universitario de Málaga, Av. de Carlos Haya, 84, E-29010 Málaga, Spain.

***Abstract*—The human brain performs a series of complex processing steps to translate spoken language from mere sound into meaning. In this article, through the analysis of electroencephalogram (EEG) signals, we try to identify those areas in which there is a substantial difference in brain activity when different languages are sought aiming at the detection of language comprehension. For this purpose, a set of sentences, both in their native and a foreign language, have been presented to various individuals. Following previous works, we defined an analysis window to determinate whether the language presented is understood or not. Within the analysis window, in the first phase, a couple of evoked potentials (ERPs) are expected, showing a positive discrimination of language, followed by fluctuations of ERPs in a second-phase window depending on subject intelligibility of the language presented.**

I. Introduction

We are currently in a globalized world, in which we are constantly exposed, not only to our native language, but also to foreign languages. Speech is just one of the many ways that language can be presented.

Thanks to the possibilities provided by the new technologies, the interest on speech detection and recognition, as well as language discrimination have been on the rise through last decade [1]. However, no definitive answer has been agreed upon how the brain reacts to different languages. Electroencephalography stands as a non-invasive technique for recording electrical activity on the surface of the scalp for research in this context.

ERPs are frequently used to study spoken-language processing since they have an excellent temporal resolution. However, not all cognitive processes have an ERP signature: only if the neurons all are oriented in a similar direction, and all receives the same stimulus, their dipoles will sum up and may be measured [2] [3].

Humans are able to process spoken-language easily, which is really surprising given the short time they have to process it. In order to determinate whether a person is able to understand a certain language, we aim to find a way to determinate this intelligibility through the analysis of raw EEG signals.

We expect to obtain differences between the ERPs from subjects when they are exposed to their native language versus a foreign one compatible with the findings in [4] in a different communication context. As suggested in [5], we chose two languages as different as possible: Spanish and Korean. We measured, and then analyzed the EEG signals obtained while different subjects were listening to these two languages.

This manuscript is organized as follows: In Sec. II, the methodology is presented, describing the participants, experiment and equipment employed. In Sec. III, the processing stages applied to the data are specified. Next, Sec. IV presents the results extracted from the analysis. Finally, conclusions are drawn in Sec. V.

II. Methods

In this section, the methodology followed to record EEG signals is described, including the description of the subjects participating, the auditory stimuli they were presented, and the specific equipment and recording procedure employed.

A. Participants

We recorded EEG data from an heterogeneous group of 9 participants (Table I). Most of them were members of Universidad de Málaga (UMA), who received no monetary compensation for their collaboration. 8 of the participants were Spanish native speakers and had no knowledge of Korean at all. Due to the complexity of finding subjects whose native language is Korean and who do not know any Spanish, our 9th subject is our single Korean native speaker. Even though, these two language have been chosen because there are only two opposite possible states: either they understand the language presented, or they don't, so non intermediate states of intelligibility can missdraw the results. This diverse linguistic competence was estimated by a survey they were asked to fill before carrying out the experiment along with an informed consent beforehand.

The experimental protocol under which the data were recorded was approved by the Comité Ético de Experimentación de la Universidad de Málaga, Reg. CEUMA: 61-2021-H.

Id	Sex	Age	Native Language	Max kΩ (EEG channel)
S_1	Male	24	Spanish	10
S_2	Male	26	Spanish	14
S_3	Male	23	Spanish	9
S_9	Male	26	Spanish	7
S_{14}	Female	21	Spanish	9
S_{15}	Male	22	Spanish	9
S_{16}	Female	21	Spanish	9
S_{17}	Female	25	Spanish	8
S_{19}	Female	21	Korean	8

TABLE I
Participants' data, including maximum measured impedance across all the EEG channels acquired for each of them.

B. Experiments

In this experiment, the subject is presented with blocks of 5 sentences of the same language in a row from among the entire list of 150 sentences for all languages (German, Korean, Spanish, English and Italian). The order in which languages are presented throughout the experiment, as much as the sentences presented in each block, is randomized. For each phrase, a different mark has been sent for synchronization.

The duration of each of these phrases is largely dictated by the language itself, with an average duration of 2.3 and 3.5 seconds for Korean and Spanish, respectively. However, for our study, we only selected the first second of each sentence for the two languages chosen, which is of interest for intelligibility detection.

Once the audio excerpt has been listened, the subject is asked if he/she understood the sentence by pressing: key 1, if they did; key 2, if they barely did; or key 3, if they didn't. The resting period is set by the subject itself.

C. Equipment

The BrainVision actiCHamp system [6] was used to capture brain activity along with actiCAP Slim active electrode system [7] for data acquisition, both the 64-channel model. The signals obtained by means of the electrodes and sensors are amplified, digitized and transferred for later visualization and storage.

Two additional electrodes are placed to measure and record optical activity (electro-oculogram or EOG): the first electrode is placed on the external part of the right eye, above the cheekbone to record the vertical eye movement (VEOG channel), and the second one is placed on the lower part of the left eye in order to record the horizontal eye movement (HEOG channel). The maximum measured impedance of each EEG channel across all the subjects is under 14 kΩ (Table I).

With regard to data acquisition, the whole experiment was designed and conducted using E-Prime software [8] for stimuli presentation and sending of synchronization marks. The data were captured using Brain Vision Recorder software [9].

III. Data analysis

In this section, the specific processing and analysis stages followed in this work are described. For this, MATLAB [10] has been used along with the dedicated EEG Fieldtrip toolbox [11] for EEG and MEG data analysis.

A. Signal Pre-processing

As mentioned before, we aim to work with raw EEG data ($S_i[n]$, with i=Id subjects). Since we are only interested in some trials, more specifically Spanish and Korean trials, we select a few of them, particularly $N_t = 5$ trials of each ($x_t[n]$, with t=1...5). Non pre-processing steps was taken beyond a high-pass FIR filter with a cut-off frequency of 0.5 Hz in order to remove continuous component of each subject ($x_{f,t}[n]$). For each subject, we averaged our trials by electrodes on different structures ($X_i[n]$), so we have differentiated Spanish and Korean averaged trials for each of them obtained according to

$$X_i[n] = \frac{1}{N_t}\sum_{t=1}^{N_t} x_{f,t}[n] \quad (1)$$

In previous works [12], a three-phase window analysis is defined: Phase 1 (100–300 ms) represents the time window in which the initial syntactic structure is formed on the basis of information about the word category; phase 2 (300–500 ms), lexical-semantic and morphosyntactic processes take place with the goal of thematic role assignment; phase 3 (500–1000 ms), the different types of information are integrated. Following this idea, we define a two-phase window analysis: phase 1 (100-500 ms), where basic processing steps are taken, such an automatic and positive recognition of the language; and phase 2 (500-1000 ms) for a more deep process of language if concepts were understood.

B. Signal Processing

However, since the task is a primarily auditory-processing, we expect to see differential activity in some specific electrodes placed in the frontal area [5]. So, we selected the electrodes: AF3, AF4, AF7, AF8, AFz, F1, F2, F3, F4, F5, F7, FC1, FC3, FC5, Fp1, Fp2 and Fz. The location and naming of the electrodes was based on the ten-twenty electrode system [13].

Recognizing a differential pattern on those electrodes was expected as some of them correspond to Brodmann's areas 11, 12 and 15 (related to decision making) since subjects are trying to determinate if the language presented seems familiar for them; and 44 and 45 areas (responsible of speech-generating) [14]. In fact, those two last areas conform the Broca's area, which concerns language comprehension when working along with Wernicke's Area [15]. It has been proved that it's activation is triggered by the factor syntactic memory but not by complexity [12], resulting in a more clearly activation for a native language.

In order to soften the high fluctuations of the waveform, and following previous researches [16], a moving average using a 150-sample window has been applied ($X_{ma}[n]$).

Once our signals have been smoothed, another average, now only for those electrodes of interest, is performed over the previous average per user, followed by the grand-average across our $N_s = 8$ Spanish speakers only ($X_g[n]$)(Fig. 1).

No matter how much we try to keep the conditions of the experiment stable, each subject will present different levels, and ERPs at slightly different times. Therefore, a common practice is to average and normalize several of these trials so that small differences cancel each other out. [17].

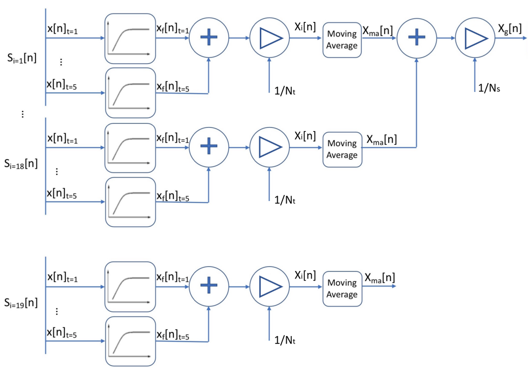
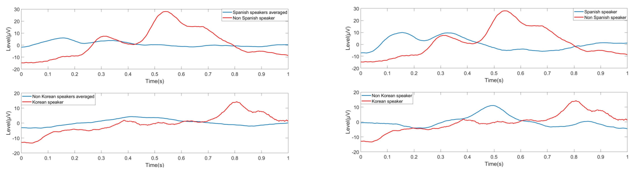


Fig. 1. Diagram of the data processing steps for each language: high-pass filter, averaging of trials and a moving average followed by a final grand-average stage of Spanish speakers.

However, because of this grand-average, we smoothed Spanish speakers signals (Fig. 2(a)) since every subject has his own level and they are not perfectly synchronized. That's the main reason why we are not able to perform direct amplitude comparisons. If we now compare, not with the average, but with a single Spanish speaker subject (Fig. 2(b)), we can see there is not such a great difference of amplitude between them. In order to make fair comparisons, the signals ($x_n = \bar{S}_{Spanish} \quad and \quad S_{19}$) are normalized (2) after averaging so that we can eliminate the differences in amplitudes but maintain the relationship between them using:

$$x_n[n] = \frac{x[n] - min(x[n])}{max(x[n]) - min(x[n])} \tag{2}$$



(a) Averaged $\bar{S}_{Spanish}$ vs. S_{19} (b) S_2 vs. S_{19}

Fig. 2. The upper figures show ERP levels when the language presented is Spanish. The lower figures show ERP levels when the language presented is Korean. Figures (a) show the amplitude difference between Spanish averaged subjects versus the Korean subject. Figures (b) show the levels of a single Spanish subject versus the Korean subject.

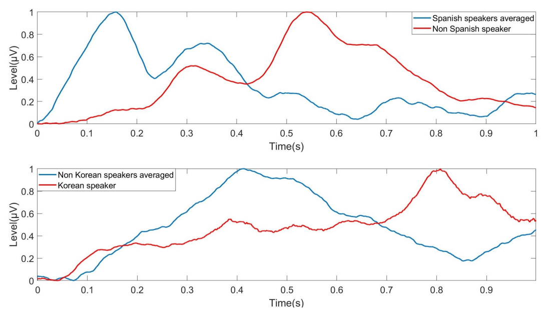


Fig. 3. Normalized ERP comparison: $\bar{S}_{Spanish}$ vs. S_{19}.

IV. Results

In this section, the results observed within this work are described. The focus was on obtaining evidence of how the brain reacts to different languages through EEG data analysis.

A. First-Phase Window: ERPs delayed

When studying the normalized ERPs (Fig. 3), we can see that, both in Spanish and Korean, there is a biphasic ERPs pattern in the first window of analysis which is clearly more defined for Spanish case. As mentioned, previous studies linked this interval (100-500 ms) to a primary process where language is recognized and analyzed at a basis morphosyntactic level [12]. This seems reasonable within our research. If we look for the exact instants these peaks occur (Table II), this pattern is clearly delayed for the foreign language case compared with the native one.This might be an indicator of how each brain reacts differently depending on the first language of speakers, showing a faster process when trying to identify and determinate the understanding they have about their own language.

	Native Speaker		Foreign Speaker	
	ERP#1	ERP#2	ERP#1	ERP#2
Spanish	155.6	328.0	309.6	540.4
Korean	160.4	380.0	236.0	432.4

TABLE II
PARTICIPANT'S ERP PEAK TIME INSTANTS (MS)

However, when measuring the difference between Spanish and Korean speakers, the Korean subject signal appears delayed around 150-200 ms with respect to Spanish speakers when the language presented is Spanish, but Spanish speakers are delayed about 50-100 ms with respect to the Korean subject when Korean is the language presented.

If we look at the surveys, our Korean subject affirmed having some knowledge about Spanish, unlike none of our Spanish subjects for the opposite case. Maybe this is reflected on those times, meaning that our Korean subject tries to stay focused on barely understanding Spanish, but since any Spanish speaker is familiarized with Korean, they gave up on trying, so they are less delayed regarding to the Korean one.

Previous researches observed this biphasic pattern, consisting of a LAN (negative deflection peaks between 300-500 ms) followed by a P600 (positive deflection at 500 ms with auditory stimuli) when native speakers are presented with morphosyntactic violations [3]. Actually, in our measures, when subjects are presented a foreign language, they present similar deflection times than when native ones are presented with this violations previously mentioned.

A similar effect is observed when comparing first with second languages between monolingual versus bilingual subjects (infants), where only monolinguals elicited an earlier response to their native language when comparing to a foreign one [18].

B. Second-Phase Window: ERPs from language process

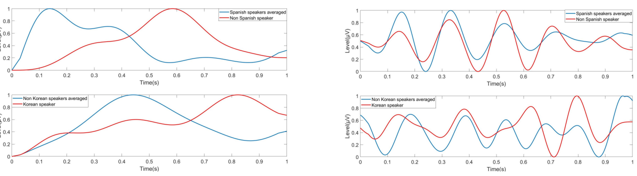
Following the model proposed in [12], at this later window we expect to see a more dense activity when subjects are able to understand the language that is being played, and this is what our measures show. When the language presented is Spanish, we can see how only Spanish speakers presents some activity after 500 ms, while for Korean subject this activity decreases. A similar behavior is observed in the opposite case, with only the Korean subject showing new higher peaks of activity after this 500 ms barrier.

C. EEG Band Based Analysis

There is no agreement on how the brain reacts to a certain language; not only when dealing with raw data, but also when working with specifics bands. Delta band (0.5-4 Hz) is highly important since encompasses the prosodic information, which is known to be important for speech intelligibility [18].

According to other researchers, Theta band (4-8 Hz) has been directly involved with attention, and some studies have related increases in Theta band to the perception of native/non-native language [19]. When an activity requires a huge amount of attention, it would be reflected with a larger amplitude. Some researches claim that been exposed to your native language seems natural to your brain, so it wouldn't require much of attention, so smaller amplitudes on this band would be expected when compared to a foreign language [20]. With an opposite point of view, some researches claim that as we are able to understand the language presented, our attention will be drawn to it unconsciously, leading to higher level of signals when facing a familiar language [21].

In order to analyze this, we apply two additional band-filtering phases to extract Delta (Fig. 4(a)) and Theta (Fig. 4(b)) band separately, with the same electrodes studied before. At the sight of previous bibliography, Delta band presents the same biphasic pattern as working with quasi-raw data. It seems coherent to have this pattern, when native speakers are exposed to audio excerpts in their first language, and delayed otherwise. Regarding to Theta band, we find it not possible to confirm one theory or another since subjects aren't precisely synchronized, so lobes are poorly formed (Fig. 4).



(a) Delta Band (b) Theta Band

Fig. 4. ERP comparison: averaged Spanish speakers vs. single Korean speaker in different frequency bands.

V. Conclusions

The current study focused on detecting language comprehension. We aim to work with quasi-raw data, so only a high-pass filter was applied as pre-processing. The data were then cut in one second long excerpts and selected depending on the language of each trial. Then it was smoothed with a moving average, followed by a grand-average only for Spanish speakers.

Our hypothesis is that native speakers shows a positive discrimination and faster processing of their language, compared to a foreign language, followed by a more deep process and analysis only if the subject is able to understand the language. This seems supported by the EEG data analysis performed.

When native speakers are exposed to their first language, a biphasic pattern composed by LAN and P600 components is drawn at the early window, revealing this positive recognition of the language and the first steps of a basic analysis of the syntactic structure. In the later window, new ERPs show deeper analysis since the language has been understood. This behavior was actually detected for both Korean and Spanish speakers in their own first language. When they face a foreign language, there is a clear delay of this biphasic patter, and there is not that much of activity in the later window since they aren't able to process any more information.

These two observations seem to be supported by the analysis of Delta and Theta frequency bands. Delta band could be then interpreted as an indicator of the activity each subject carries out to understand the language they are exposed and Theta as an indicator of the effort required to do so. Native speakers only have constant delta activity in their first language, and the level they show at Theta band seems directly related to it. However, since the Korean subject has a basic knowledge of Spanish, it could be interpreted the activity in both bands, as much as the delay in some component, as a truly attempt to keep on understanding it. Nevertheless, further experiments with an enlarged data set should be carried out in order to extract more rigorous statistical information so we can develop some classifiers based on Neural Networks.

Acknowledgements

This work was funded by Junta de Andalucía: Proyectos de I+D+i en el ámbito del Plan Andaluz de Investigación, Desarrollo e Innovación (PAIDI 2020), under Project No. PY20_00237, and Programa Operativo FEDER Andalucía 2014-2020, under Project No. UMA18-FEDERJA-023. It has been possible under a grant from the program "Ayuda de Iniciación y Transferencia", Universidad de Málaga. It has been done within Campus de Excelencia Internacional Andalucía Tech.

References

[1] M. Koctúrová, and J. Juhár, "A Novel Approach to EEG Speech Activity Detection with Visual Stimuli and Mobile BCI," *Applied Sciences*, vol. 11, no. 2, pp. 674, 2021.
[2] S. J. Luck, "An introduction to the Event-Related Potential Technique," *A Bradford Book, MIT Press*, 2014.
[3] M. Pélessier, "Comparing ERPs between native speakers and second language learners: Dealing with individual variability," In Interpreting language-learning data *Berlin: Language Science Press*, pp. 39-69, 2020.
[4] A. Pérez, G. Dumas, M. Karadag and J.A. Duñabeitia. "Differential brain-to-brain entrainment while speaking and listening in native and foreign languages," *Cortex*, vol. 111, pp. 303-315, 2019.
[5] A. Soman, C. R. Madhavan, K. Sarkar and S. Ganapathyr, "An EEG study on the brain representations in language learning," *Biomedical Physics and Engineering Express*, vol. 5, no. 2, 2019.
[6] actiCHamp (64 channels) [Apparatus], *Brain Products GmbH*, 2016.
[7] actiCAP (64 channels) [Apparatus], *Brain Products GmbH*, 2015.
[8] W. Schneider, A. Eschman and A. Zuccolotto, "E-Prime User's Guide," *Pittsburgh: Psychology Software Tools Inc.*, 2012.
[9] BrainVision Recorder (Vers. 1.23.001) [Software], *Brain Products GmbH*, 2020.
[10] MATLAB (R2021b) [Software], *Mathworks*, 2021.
[11] FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data [Software] *Computational Intelligence and Neuroscience*, 2011.
[12] A. D. Friederici, "Towards a neural basis of auditory sentence processing," *Trends in cognitive sciences*, vol. 6, no. 2, pp. 78-84, 2002.
[13] H. H. Jasper, "The ten-twenty electrode system of the International Federation," *Electroencephalography and clinical neurophysiology*, vol. 10, pp. 367-380, 1958.
[14] F. R. Lapuente, M. J. Rabadán and M. del Pino Sánchez, "Neuropsicología," *Diego Marín Librero Editor*, 2010.
[15] J. Moini and P. Piran, "Functional and Clinical Neuroanatomy. A Guide for Health Care Professionals," *Academic Press*, 2020.
[16] A. Kawala-Sterniuk, M. Podpora, M. Pelc, M. Blaszczyszyn, E.J. Gorzelanczyk, R. Martinek and S. Ozana, "Comparison of Smoothing Filters in Analysis of EEG Data for the Medical Diagnostics Purposes," *Sensors*, vol. 20, no. 3, pp. 807, 2020.
[17] S. Stober, A. Sternin, A. M. Owen and J. A. Grahn, "Deep Feature Learning for EEG recordings," *The Brain and Mind Institute. University of Western Ontario*, 2016.
[18] J. Vanthornhout, L. Decruy, J. Wouters, J. Z. Simon and T. Francart, "Speech Intelligibility Predicted from Neural Entrainment of the Speech Envelope," *Journal of the Association for Research in Otolaryngology: JARO*, vol. 19, no. 2, pp. 181-191, 2018.
[19] L. N. Garcia, C. Guerrero-Mosquera, M. Colomer, and N. Sebastian-Galles, "Evoked and oscillatory EEG activity differentiates language discrimination in young monolingual and bilingual infants," *Scientific Reports*, vol. 8, no. 2770, 2018.
[20] A. N. Bosseler, S. Taulu, E. Pihko, J. P. Mäkelä, T. Imada, A. Ahone and P. K. Kuhl, "Theta brain rhythms index perceptual narrowing in infant speech perception," *Frontiers in Psychology*, vol. 4, no. 690, 2013.
[21] A. Pérez, M. Carreiras, M. G. Dowens, and J. A. Duñabeitia, "Differential oscillatory encoding of foreign speech," *Elsevier Brain and Language*, vol. 147, pp. 51-57, 2015.