Neural correlates and L2 lexical stress learning: an fMRI study

Schwab, Sandra ; Mouthon, Michael ; Salvadori, Justine ; Silva, Eugenia Ferreira da ; Yakoub, Ilona ; Giroud, Nathalie ; Annoni, Jean-Marie

DOI: https://doi.org/10.21437/speechprosody.2022-167

Originally published at:
Schwab, Sandra; Mouthon, Michael; Salvadori, Justine; Silva, Eugenia Ferreira da; Yakoub, Ilona; Giroud, Nathalie; Annoni, Jean-Marie (2022). Neural correlates and L2 lexical stress learning: an fMRI study. In: Speech Prosody 2022, Lisbon, Portugal, 23 May 2022 - 26 May 2022, Speech Prosody Special Interest Group (SProSIG).
DOI: https://doi.org/10.21437/speechprosody.2022-167

Speech Prosody 2022
23-26 May 2022, Lisbon, Portugal

Neural correlates and L2 lexical stress learning: an fMRI study

Sandra Schwab[1,2], *Michael Mouthon*[1], *Justine Salvadori*[1], *Eugenia Ferreira da Silva*[1], *Ilona Yakoub*[3], *Nathalie Giroud*[3] *& Jean-Marie Annoni*[1]

[1]University of Fribourg, Switzerland, [2]University of Bern, Switzerland, [3]University of Zurich, Switzerland

[sandra.schwab; michael.mouthon; Justine.salvadori; jean-marie.annoni; eugenia.ferreiradasilva]@unifr.ch, [ilona.yakoub; Nathalie.giroud]@uzh.ch

Abstract

Stress detection in a second/foreign language (L2) is a complex task which depends of different linguistic and cognitive factors. The objective was to investigate to what extent the degree of activation in specific brain regions involved in L2 stress perception was related to the listeners' ability to detect L2 stress before training, and to their stress learning capacity after training. French-speaking participants with no knowledge of Spanish took part in an fMRI study in Spanish, as well as in a pre/post-training experiment, also in Spanish.

Results showed that native listeners of French improved their stress perception in a L2 after training, even though the training effect was rather subtle and interindividually variable. The results also revealed that there was a link between the degree of neural activation in left inferior frontal gyrus and the participants' performance in L2 stress identification before training. No significant correlation was found with the amount of learning after training. These data highlight the fact the interindividual differences observed in L2 stress processing might be (at least partially) related to neural interindividual differences.

Index Terms: lexical stress perception, neural activation, inferior frontal gyrus, training effect

1. Introduction

The discrimination of lexical stress contrasts in a foreign/second language (hereafter 'L2'[1]) (e.g., English import versus import; Spanish número versus numero)[2] constitutes a complicated task for some listeners, especially for native speakers of languages with predictable lexical stress (e.g., French). The listeners' ability to detect L2 stress depends not only on their native language (L1) and its stress properties (fixed versus free stress, default stress pattern, stress acoustic cues; [1], [2], [3]), but also on cognitive factors (e.g., working memory or phonological awareness; [4]) and the learners' musical aptitude ([5]). Recent research has shown that, despite this so-called stress 'deafness', learners were able to overcome (part of) their L2 stress detection difficulties thanks to a perceptual training. Using a word-shape matching task in their 4-hour perceptual training, [6] showed an approx. 10% improvement after training in different perceptual tasks, such as stress identification and Odd-One-out. However, there was a certain degree of interindividual variability in the stress detection ability before training, as well as in the amount of improvement after training.

The present research explores to what extent neural correlates of L2 stress processing (before training) are related with the listeners' L2 stress identification performance before training, as well as with the amount of improvement after training. Neural correlates underlying word stress processing are still not well understood. Although most researchers agree on the strong lateralization of activation to the left hemisphere during stress processing in the listeners' L1, specially in fronto-temporal areas ([7]), some found bilateral activation ([8], [9]). To our knowledge, similar fMRI studies examining L2 stress perception do not exist. Nevertheless, of particular interest for the present project is the research carried out by [10] on tone perception in a second language. They showed a relationship between neural correlates and language performance on an individual level, not only after training, but, interestingly, also before training. 'Successful' learners (i.e., participants who learned more from training) showed before training increased activation mainly in bilateral superior/middle temporal regions, whereas the less successful learners showed increased activation in right middle frontal areas.

2. Objectives

This study examined L2 stress acquisition by means of neuroimaging and behavioral experiments. We conducted an fMRI experiment in L2 following the methodology used in L1 by [9]. The fMRI experiment was carried out once, before the behavioral experiment, hence before training. The behavioral experiment served to assess the listeners' ability to identify L2 stress before training, as well as their L2 stress learning capacity after training.

The objective was to investigate to what extent the degree of activation in specific brain regions involved in L2 stress perception before training was related to the listeners' ability to detect L2 stress before training, and to their stress learning capacity (i.e., amount of improvement after training). We hypothesized that L2 stress processing involves fronto-temporal areas as in L1. However, although literature on L2 induced activation suggests convergent anatomico-functional relationship between L2 and L1 (e.g., [11], [12]) given the lack of evidence on this specific process, we were not able to make

[1] 'Second language' or 'L2' does not necessarily correspond to the second language that participants have learned. It is used in this paper as a synonym for 'foreign language'.

[2] The underlined syllable in these examples, and in the rest of the paper, corresponds to the stressed syllable.

 10.21437/SpeechProsody.2022-167

any claim about the lateralization of L2 stress processing. Then, we predicted that L2 listeners who show large activation in the fronto-temporal regions (whether left lateralized or bilateral) experience more stress perception difficulties and less improvement after training than L2 listeners with little activation in these areas.

3. Methods

3.1. Participants

Thirty students participated in the experiment (18 women; range = 18-27 years; mean age = 22,67; st. dev. = 2.37). They were all recruited in Switzerland at University of Fribourg or Haute Ecole in Fribourg via announcements on social media. Participants were native speakers of Swiss or standard French. They had no knowledge of Spanish, Italian, or Portuguese (i.e., free-stress romance languages). Since German and English are mandatory disciplines in the Swiss educational system, all participants had school knowledge of these two languages. Although thirteen participants had received musical training (among which 7 were still playing an instrument at the moment of the experiment), none of them was a professional musician. Participants were paid for their participation. The study received ethics approval from the Ethics Committee of the Psychology Department of the University of Fribourg (IRB no 530).

3.2. Experimental design

The experimental design was composed of three phases (see Figure 1). In Phase 1, participants performed an fMRI experiment and a pretest (i.e., identification task). In Phase 2, they received a 4-hour training in Spanish. In Phase 3, they performed a posttest which consisted again in an identification task.

Running behavioral experiments in Phases 1 and 3 (i.e., before and after training) allowed us to examine the training effect (i.e., improvement after training). Conducting the fMRI experiment in Phase 1 allowed us to relate neural correlates before training to the pre-training behavioral performance and to the behavioral training effect.

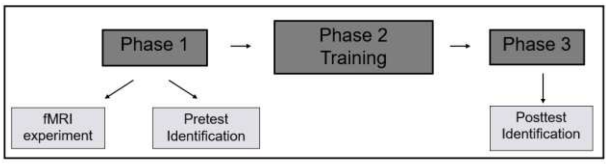

Figure 1: *Experimental design.*

3.3. fMRI paradigm

3.3.1. Task and material

In the fMRI experiment, the participants performed a discrimination task modeled on that used in [9]. Participants were presented with pairs of auditory trisyllabic Spanish words produced by a female native Speaker of Castilian Spanish. They were asked to indicate whether the two words were identical or different. In the 'different' pairs, items differed either in the quality of the final vowel (e.g., valoro versus valore) or in the position of word stress (e.g., valoro versus valoró). Both 'vowel' and 'stress' different pairs showed phonological contrasts, the former at the segmental level, the latter at the suprasegmental level. In the 'vowel' different pairs, one of the items was always the 1sg present indicative of the verb (e.g., valoro, en. I value) and the other item the 1sg/3sg present subjunctive of the same verb (e.g., valore, en. that I/he/she value). The 'stress' different pairs were always composed of the 1sg present indicative of the verb (e.g., valoro, en. I value) and 3sg simple past tense of the same verb (e.g., valoró, eng. he/she valued). Since the exact same verbs were used in the 'vowel' and 'stress' conditions, the lexical frequency of the verbs did not need to be controlled.

3.3.2. Data acquisition and pre-processing

The 'vowel' condition served as a control condition to be compared with the experimental 'stress' condition in order to identify the specific regions activated during stress processing. The two conditions were presented using a block design. MRI data collection included two fMRI runs of 9 minutes that were acquired successively with a 3T GE scanner (acquisition details: TR = 2000 ms, TE = 30 ms, flip angle = 85°, voxel size = 2.3 x 2.3 x 3 mm).

MRI data preprocessing and analyses were conducted with SPM12 (UCL, London). Functional images were preprocessed with a standard pipeline (slice timing, spatial realignment and unwarping, co-registration of anatomic scan on fMRI, Unified-segmentation of anatomic scan, normalization, smoothing). The resulting images were analyzed at the individual subject level using a general linear model (GLM). fMRI signal was modelized as condition-specific block of 22.63s of duration convolved with the hemodynamic response function (HRF). A high-pass filter with a 1/128 Hz threshold was applied at time series to remove low frequency noise and signal drifts and an autoregressive function (AR(1)) was implemented to correct temporal correlations between neighboring voxels.

3.3.3. Data analysis

The contrast between Stress and Vowel block was sent to a one-sample t-test to study the general difference of brain activity between these two conditions. The results were studied on the whole brain space with the statistical threshold Family-Wise Error corrected for multiple comparison at the peak level ($pFWE<0.05$) as well as at cluster level (minimal size of 11 significant contiguous voxels which corresponds to $pFDR<0.05$). Further analyses were performed only in areas which showed significant differences between Stress and Vowel conditions (i.e., regions of interest).

In continuation, the means of neural activation for the regions of interest for each participant are expressed by numeric values resulting from the contrast between Stress and Vowel conditions. The higher the value, the larger the difference between the activation for Stress and Vowel processing.

3.4. Behavioral experiments

3.4.1. Pre-/posttests

In the behavioral pre- and posttests, participants performed an identification task in which they heard trisyllabic Spanish words (e.g., titulo) and had to indicate which syllable was stressed. The pretest was composed of 120 stimuli, while the posttest comprised 180 stimuli (120 stimuli from pretest + 60 new stimuli). All words were produced by two Spanish speakers with a falling or a rising intonation (see [1] for details). The percent correct for pre- and posttest was collected for each participant.

3.4.2. Training

Participants received a 4-hour training on Spanish accentuation divided into eight 30-min sessions over two weeks (see [1] for details). The training approximated a situation of immersion, where the participants received no explicit explanations about accentuation in Spanish and had to perform only one task during the entire training, namely, a Word/Shape Matching task. Listeners heard a word and 4 shapes (i.e., similar to Tetris shapes) in different colors appeared on the screen. They had to click on the shape they thought corresponded to the word they heard. After giving their response, they received feedback: they heard the word again and only the correct shape stayed on the screen. The feedback enabled the listeners to learn the association between the words and the shapes. The outcome measures of the training were not further analyzed in the present research.

3.5. Data analysis

A paired t-test was first used to examine the training effect (i.e., difference between post- and pretest). Then, correlation analyses were run to determine the relationship between the degree of neural activation in the specific regions of interest and 1) percent correct at pre-tests; 2) training effect (i.e., difference between pre- and post-tests).

4. Results

4.1. Training effect

As shown in Figure 2, we observe a progression from pre- (64.52.10%, stdev = 14.54) to posttests (73.91%, stdev = 16.88; $t(29) = 6.18$; $p < .001$; Cohen's $d = 1.13$). The training effect (i.e., difference between post- and pretest) was thus 9.3%. As expected, we note however a certain amount of variability in the training effect among participants (see Figure 3).

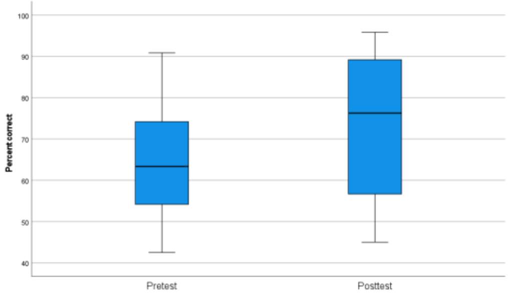

Figure 2: *Identification percent correct for pre- and posttest*

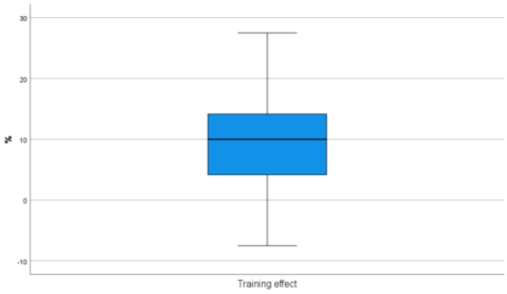

Figure 3: *Training effect (i.e., difference between post- and pretest).*

4.2. Neural activation in bilateral inferior frontal gyrus (IFG)

fMRI whole brain analyses showed differences of brain activation between Stress and Vowel conditions in the bilateral inferior frontal gyrus (IFG) and in right middle/superior temporal gyrus (pFWE<0.05, min cluster size=11). We focused thus the subsequent analyses on these particular areas.

Figure 4 presents the distribution of the activation difference between Stress and Vowel conditions in the left and right IFG and in the middle/superior temporal gyrus. As mentioned earlier, the higher the value the larger the activation difference between Stress and Vowel conditions. As can be seen, the neural activation in the three brain regions was different from 0 despite the presence of a certain variability among participants.

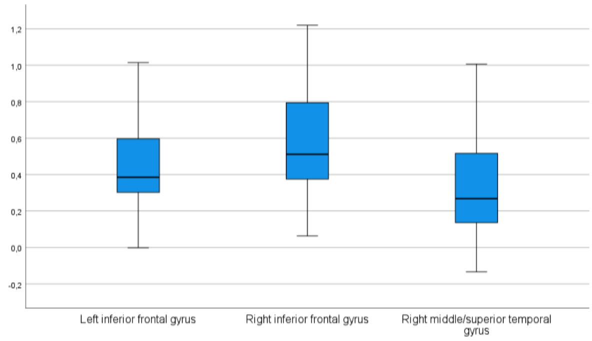

Figure 4: *Activation difference between Stress and Vowel conditions in left and right inferior frontal gyrus, and in right middle/superior temporal gyrus.*

In the continuation of the paper, the "activation difference between Stress and Vowel conditions" will be referred to as "neural activation" for the sake of simplification.

4.3. Relationship between neural activation and L2 stress processing

Figures 5 and 6 show the identification percent correct before training and the training effect (i.e., difference between post- and pretests), respectively, as a function of the neural activation in left and right IFG and in right middle/superior temporal gyrus.

As can be seen in Figure 5, we observe a marginal negative correlation between the listeners' performance before training (i.e., baseline performance) and the neural activation in the left IFG ($r = -.34$, $n = 30$, $p = .07$), but no significant correlation in the right IFG ($r = -.14$, $n = 30$, $p = .48$), nor in the right middle/superior temporal gyrus ($r = -.0004$, $n = 30$, $p = .98$). The participants with lower stress identification performance (before training) tended to present a larger neural activation in left IFG than participants with high performance. Such data seem to suggest that the larger the difficulties in identifying stress position in the behavioral task, the larger the activation for stress processing in left IFG. This relationship could not be observed for the right IFG nor for the right middle/superior temporal gyrus.

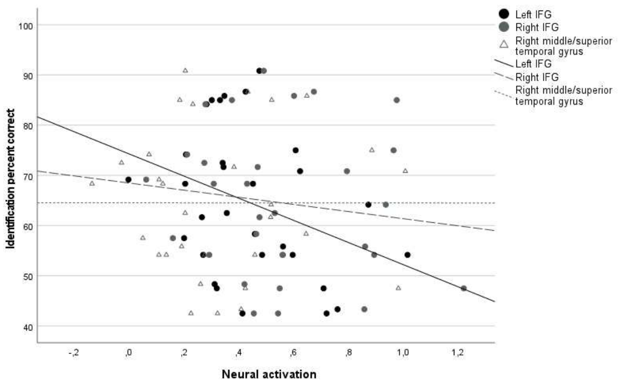

Figure 5: *Identification percent correct before training as a function of neural activation in left and right inferior frontal gyrus, and in the right middle/superior temporal gyrus.*

Figure 6 indicates no significant relationship between the listeners' training effect (i.e., difference between post- and pretest) and prior neural activation in left IFG (r = .04, n = 30, p = .84), right IFG (r = .01, n = 30, p = .98) or middle/superior temporal gyrus (r = .17, n = 30, p = .36). In other words, the amount of extra cognitive resources for stress processing did not seem to be related to the amount of learning after training.

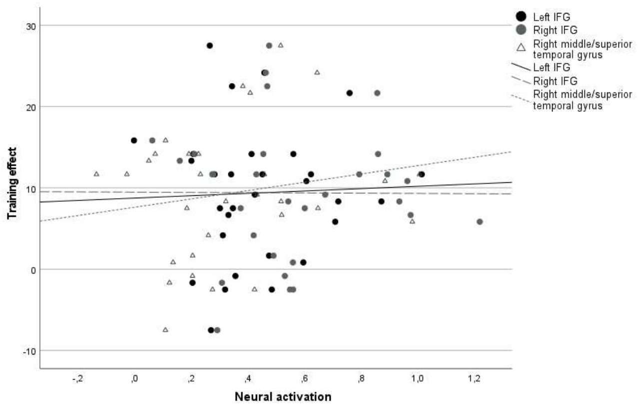

Figure 6: *Training effect as a function of neural activation in left and right inferior frontal gyrus and middle/superior temporal gyrus*

5. Discussion

Our results first showed that French-speaking participants did improve their identification of stress position in Spanish after training, Noteworthily, the average training effect (9.38%) was similar to the one obtained by [1] with the same methodology (11.83%). As expected, a certain degree of variability was observed, not only in the participants' performance before training, but also in the amount of improvement after training.

Second, our results revealed that L2 stress processing, in comparison with vowel processing, mainly involved the brain areas located in the bilateral inferior frontal gyrus as well as in right middle/superior temporal gyrus. Since we found stronger activation for L2 stress processing in frontal regions and (only) right temporal area, our finding is in partial agreement with [9] who showed bilateral activation of fronto-temporal regions for L1 stress processing in German pseudowords. They are however in agreement with data showing that decreased comprehension expertise in L2 is associated with lower brain activity (e.g., [13]). Since our study is the first attempt to identify brain regions involved in L2 stress processing, further research is still needed to confirm our conclusions.

Third, our findings highlighted a marginal relation between the listeners' baseline performance and the neural activation in left IFG, but not in the right IFG nor in the right middle/superior temporal gyrus. The less neural resources were needed for stress processing (i.e., the smaller the left IFG activation), the better the participants seem to be at identifying stress position before training. This finding partially confirms our hypothesis that L2 listeners with large activation in fronto-temporal regions experience more stress perception difficulties than L2 listeners with little activation in these areas, since we found the presence of such a link for the (left) inferior frontal regions, but not for temporal areas.

Finally, no significant relationship was observed between neural activation and the amount of improvement after training. This result is not in agreement with [10] who found such a relationship for L2 tone perception. One possible explanation resides in the nature of our training. A 4-hour training might not be sufficient for French-speaking listeners to overcome their perceptual difficulties with stress detection, as suggested by the slight training effect (less than 10%). As a consequence, neural correlates might not be able capture such limited amount of learning. This explanation is in line with [14], who found, using the same pre/post-training design, a relationship between P3b amplitude and the listeners' performance before training, but not with their improvement score (i.e., training effect). Thus, although our research could not evidence the presence of a link between neural activation and L2 stress learning, it is plausible to consider that such a link could be pointed out using another methodology (i.e., longer and more efficient training)

To conclude it is worth mentioning that the fMRI experiment has been carried out only before training. The absence of fMRI measurements after training makes it thus impossible to observe possible training-related neural changes, such as those observed by [10] for L2 tone learning.

6. Conclusions

Our findings have important implications for the acquisition of L2 prosody. We have shown first that native listeners of French were able to improve their stress perception in a L2, even though the training effect was rather subtle and interindividually variable. Our results then revealed that there was a link between neural activation and the performance in L2 stress identification, even before training. This conclusion highlights the fact the interindividual differences observed in L2 stress processing (and possibly learning) might be (at least partially) related to neural interindividual differences.

7. Acknowledgements

This research was funded by SNF Scientific Exchange program (IZSEZ0_180523 / 1), Pool recherche 2019 from University of Fribourg and Fond du Centenaire (no 797), University of Fribourg. We would like to thank Prof. Rachel Theodore (University of Connecticut, USA) for her help in designing the fMRI paradigm.

8. References

[1] E. Dupoux, C. Pallier, N. Sebastián and J. Mehler, "A destressing 'deafness' in French?," *J. Memory Lang*., 36(3), pp.406-421, 1997

[2] S. Schwab, S. and V. Dellwo, "Intonation and talker variability in the discrimination of Spanish lexical stress contrasts by Spanish, German and French listeners, " *J. Acoust. Soc. Am.*, 142 (4), pp. 2419–2429, 2017.

[3] M. Ortega-Llebaria, H. Gu and J. Fan, "English Speakers' Perception of Spanish Lexical Stress: Context-Driven L2 Stress Perception," *J. Phon*., 41, pp. 186–97, 2013.

[4] F. Honbolygó, A. Kóbor and V. Csépe, "Cognitive components of foreign word stress processing difficulty in speakers of a native language with non-contrastive stress," *International Journal of Bilingualism*, September 4, 2017. http://journals.sagepub.com/doi/abs/10.1177/1367006917728393#articleCitationDownloadContainer.

[5] P. Degrave, "Can music help learners and teachers in word stress perception?," *Travaux du Cercle Belge de Linguistique*, 11, pp. 1-20, 2017

[6] S. Schwab and V. Dellwo. "Explicit versus non-explicit prosodic training in the learning of Spanish L2 stress contrasts by French listeners," *Journal of Second Language Studies*, 2021. http://doi.org/10.1075/jsls.21017.sch

[7] A. Aleman, E. Formisano, H. Koppenhagen, P. Hagoort, E. H. F. de Haan and K. Kahn, "The functional neuroanatomy of metrical stress evaluation of perceived and imagined spoken words," *Cereb.Cortex* 15, pp. 221-228, 2005.

[8] M. Heisterueber, E. Klein, K. Willmes, D. Heim and F. Domahs, "Processing word prosody—behavioral and neuroimaging evidence for heterogeneous performance in a language with variable stress," *Frontiers in Psychology*, *5*, pp. 365, 2014. http://doi.org/10.3389/fpsyg.2014.00365

[9] E. Klein, U. Domahs, M. Grande, and F. Domahs, "Neuro-cognitive foundations of word stress processing—evidence from fMRI," *Behav.Brain Funct.* 7(1), pp. 15, 2011.

[10] P. C. M. Wong, T. K. Perrachione and T. B. Parrish, "Neural characteristics of successful and less successful speech and word learning in adults," *Human Brain Mapping*, 28(10), pp. 995–1006, 2007.

[11] D., Klein, B. Milner, R. J. Zatorre, E. Meyer and A. C.Evans, "The neural substrates underlying word generation: a bilingual functional imaging study". *Proc. Natl. Acad. Sci. USA*, 92, pp. 2899–2903, 1995.

[12] D. Perani and J. Abutalebi,"The neural basis of first and second language processing," *Curr. Opin. Neurobiol*., 15, pp. 202–206, 2005.

[13] J. Abutalebi, "Neural aspects of second language representation and language control, " *Acta Psychol*., 128, pp. 466–478, 2008.

[14] S. Schwab, N. Giroud, M. Meyer and V. Dellwo, "Working memory and not acoustic sensitivity is related to stress processing ability in a foreign language: An ERP study," *Journal of Neurolinguistics*, 55, pp. 1-15, 2020.