Effects of variability in second language learning in adults

Rebeca Sánchez López

Supervisor: Cristina Baus Márquez (CBC, UPF)

Co-supervisor: Marc Gimeno Martínez (CBC, UPF)

Key words: variability, pseudowords, signs, language and L2 learning.

Character's: 28.125

End-of-Degree Dissertation

Human Biology (2015-2020) – Pompeu Fabra University

Speech production and bilingualism – Center for Brain and Cognition

Barcelona, 22 June 2020

SUMMARY

Learning a second language is a difficult task for adults. The present thesis explored how variability in the number of teachers influences second language learning and how this process is modulated by the similarity of the language to be learned with the first language of the learner. To prove this, two experiments have been conducted: the first one using pseudowords, and the second with Catalan Sign Language, LSC. 54 native speakers of Spanish and Catalan participated in each experiment. Variability in the number of teachers was compared in three conditions: no variability— six repetitions of each word in the voice of one speaker—, moderate variability— two repetitions of each word in the voice of three different speakers—, high variability— one repetition of each word in the voice of six different speakers—. Learning accuracy was measured in two tasks: picture-to-L2 and L2-to-L1. The results in both experiments revealed a different influence of variability depending on the task. In the first experiment an effect of variability was observed in the picture-to-L2 task. In the second experiment, no effect of variability was observed in any of the tasks. Considering a possible ceiling effect, a retest of the second experiment was conducted two weeks later. The results reported an effect of variability in the picture-to-L2 task. These results show that variability enhances second language learning in different modalities regardless of the similarity of the language to be learned with the first language of the learner.

INTRODUCTION

Learning a second language (L2) is a difficult task, especially for adults. Successful outcomes of L2 learning are highly variable and depend on several factors, some related to the learner, some related to the learning exposure, and some others related to the language to be learned. Examples of those factors would be for instance the age of the learner, the acoustic variability of the learning context or the closeness between the language to be learn and the first language (Sadakata & McQueen 2014; Tagarelli et al. 2016).

The present thesis focussed on the two latter factors to explore how acoustic variability, indexed by the number of talkers, influences second language learning and how learning is modulated by the similarity of the language to be learned with the first language of the learner. To answer the first question, we replicated and expanded Barcroft and Sommers study (2005), showing that word learning improves with increased variability in the number of teachers (but see Sinkeviciute study (2019), for the opposite pattern in children). To answer the second question, we compared the learning patterns when participants were learning an oral language (invented) and a sign language (Barcroft & Sommers 2005; Giannakopoulou et al. 2017; Sinkeviciute et al. 2019).

Barcroft and Sommers (2005) examined the effect of presenting target words in an acoustically varied format using speech produced by multiple talkers during second language vocabulary learning. The experiment consisted of 60 native speakers (NSs) of English as participants with no previous instruction in Spanish, and 24 Spanish words were selected for the learning task. Each participant learned eight words in each of the three following conditions: no variability—six repetitions of each word in the voice of one speaker—, moderate variability—two repetitions of each word in the voice of three different speakers—, and high variability—one repetition of each word in the voice of six different speakers—. The learning outcomes were evaluated in two speech production tasks, which consisted in a picture-to-L2 task and L2-to-L1 translation. The results showed that accuracy was significantly better for words learned in the high variability condition for the picture-to-L2 and also for the L2-to-L1 recall task. Their conclusion was that high variability condition generated six different variants of the word form (Figure 1), which encouraged powerful lexical representations of new L2 vocabulary and which could ultimately be used as a technique for L2 vocabulary learning promotion (Barcroft & Sommers 2005).

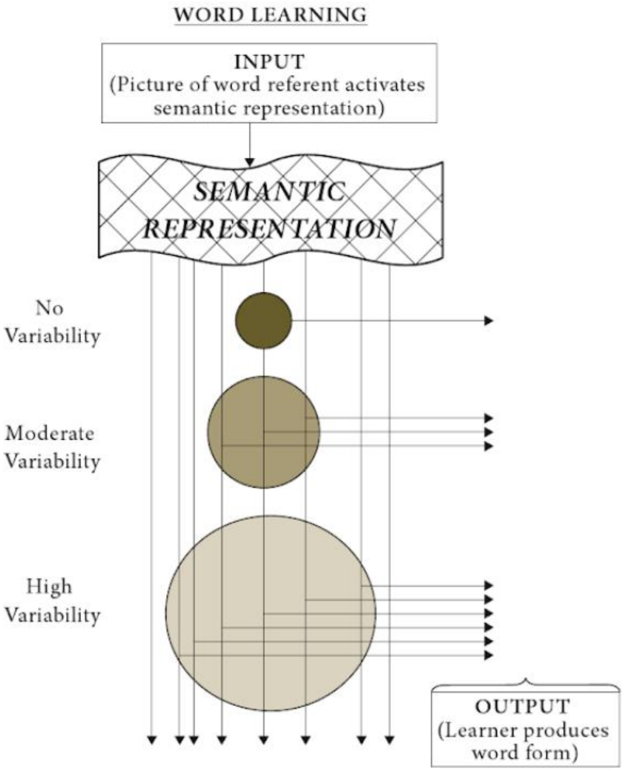

Figure 1. Model of acoustically varied input and lexical representation. Barcroft, Joe. 2015.

The purpose of our study was to replicate and expand upon previous research on speaker variability and L2 vocabulary learning to improve L2 learning in adults. To do so, two experiments have been conducted. The first study is similar to Barcroft's, but using an invented language (pseudowords), and the second one, used the same design but exploring a language in a different modality (Catalan Sign Language, LSC).

Since pseudowords and Spanish are two spoken languages, they share phonemes. For this reason, our hypothesis is that results would be more accurate with increased variability in both tasks. Besides, between-talker variability might be more important perceptually than within-talker variability, because in normal conversations talker changes often serve as a signal that a new topic or a new piece of information is about to be presented (Barcroft & Sommers 2005). Finally, talker variability has proved that, under certain conditions, produces an improved identification and retention of spoken words and non-native phonetic contrasts (Goldinger, Pisoni & Logan 1991; Lively, Logan & Pisoni 1993; Hardison 2003; Barcroft & Sommers 2005; Dong et al. 2019). Thus, talker variability might increase the strength of memory connections between individual word forms and referents to a greater degree than just one speaker (Barcroft 2001).

On the other hand, LSC is not a spoken language as Spanish is, and, therefore, they do not share phonemes, even though we expect the same results because it is also a language. Additionally, it is known from other studies that variability also improves motor learning (Dhawale, Smith & Ölveczky 2017). For this reason, our hypothesis is that variability may improve LSC learning and that accuracy results would be higher than in pseudowords experiment as LSC can be easier to learn.

Experiment 1

METHODS

Participants

A total of 54 participants, consisting of 16 males and 38 females, were recruited from the Center for Brain and Cognition subjects' database. They were rewarded for their participation. Due to the current pandemic situation, we adapted the experiment to its online form, and for this reason, the recall times (RT) could not be calculated. Participants were tested individually in front of their computers. The list used for each participant was assigned randomly. Stimuli presentation in the learning and production phases were executed with E-Prime Script, a programming software that allows to present time-controlled sequences of the pictures and their corresponding audios during exposure. We recorded the screen and shared the video with the subjects afterwards. In the learning phase, the stimuli were presented through the three conditions yet mentioned (no variability, moderate variability and high variability); the participants' role was to memorize the words using the technique that best worked for them (e.g. loud repetitions, mental association, etc). The testing phase consisted of two tasks: picture-to-L2 and L2-to-L1 recall. In these tasks, pictures and audios were presented randomly; in addition, for the L2-to-L1 recall task, voices were novel in order to remove any effect related to this.

Materials

The initial stimuli consisted of 48 Spanish words and pictures. This group of target words was formed by names which varied in a range of syllables from 1 to 4 (average 2,72). Using the program Wuggy (a pseudoword generator particularly geared towards making nonwords for psycholinguistic experiments) we generated 48 pseudowords (twice the number of target words used in Barcroft's experiment) from the initial Spanish word set.

Experimental words were divided into three categories containing 16 words each; we made sure that words did not differ in number of phonemes (average 6,06):

Word set 1: *Cecefo, minón, ina, pemalero, anlecalora, rufeso, arpel, oraka, salana, vansusta, tisbilla, hosmurcue, nafleta, jibi, leta and beceserca.*
Word set 2: *Vetruza, tisbero, suntilla, ócemo, médano, ricuento, aliza, cacebla, percel, lepón, morba, sama, cunvo, edo, angrebador and harniza.*
Word set 3: *Nívuton, mecosar, sorano, cerocho, faumante, acefo, cardetus, mafralo, lufón, jobro, crena, sible, gubra, sira, vavecoa and miza.*

Recordings were made in a soundproof room, through the software Audacity (a multi-track audio editor and recorder), by 15 different speakers (7 males and 8 females, all of them native speakers (NSs) of Spanish). Six speakers were used for the learning phase (3 males and 3 females), whereas the remaining nine speakers were used for the production tasks to ensure the use of novel voices. They were asked to record the 48 pseudowords in a neutral voice type, and each speaker produced two lists of recordings in order for the researcher to choose the best input quality (2 of the 4 different lists) besides, each audio was fragmented to a single wav file.

Procedure

The design of the experiment, including all conditions, is to be found in Appendix A. Word sets and presentation orders were counterbalanced across the learning conditions so that each participant underwent all three of them—no variability, moderate variability and high variability— but with different learning order and speakers, aimed at the outcome not being associated to a specific voice type. A table consisting of three different SETs was made with six different sublists, one per participant; for instance, the subject given the SET 1, sublist 1, listens to speaker number 1 producing word set 1 (six times each pseudoword), corresponding to the no variability condition set. Furthermore, the word set 2 represent the moderate variability condition because the subject listens to speakers 1, 4 and 5 (2 repetitions of each pseudoword per speaker); in contrast, the subject given the SET 2, sublist 1, listens to six different speakers producing every pseudoword in word set 1 (just one repetition of each pseudoword per speaker), this being the high variability condition. Six repetitions of each word for a total of 288 productions were presented to each participant. Order of presentation was constant across the six repetitions while speaker changed across sublists. Additionally, we came up with 6 more SETs to change the order of appearance of the different conditions. Thus, SET 1 starts with no variability condition, SET 4 initiates with moderate variability condition and, finally, SET 7 commences with high variability condition; however, all of them have the same word set in each condition produced by the same speakers. The same happens with SETs 2, 5 and 8 and with SETs 3, 6 and 9.

Data analysis

The criterion used to score the picture-to-L2 task and the L2-to-L1 recall was the following: subjects were given 1 point for a completely correct production and 0 points for any other answers; no answer also meant 0 points. In some cases, the participant misunderstood the image, but whether the answer was similar or a synonym for the original one (like peach and apricot, for instance) the answer was marked as 1 too. Plots with the mean accuracy by variability condition were obtained by means of the package ggplot2 and dplyr for the R computer environment. The analysis of the experiment, including all commands, is to be found in Appendix B.

The results were analysed by mixed linear models— binomial family— with task and condition (and the interaction between them) as fixed factors, while subjects and items represented random factors. The binomial family considers the accuracy as a numeric factor (Jaeger 2008). The mixed model tested the variability effect “no variability, moderate variability, high variability” in both recall tasks “Picture-to-L2” and “L2-to-L1” analysing the condition and task effect and the interaction between them. The no variability condition was selected as intercept and compared with the variability conditions (medium and high). The commands of the mixed model analysis are to be found in Appendix C.

RESULTS

Firstly, means and standard errors (SE) for accuracy based on variability were calculated as shown in Figure 2. As it can be seen in the graph, in the picture-to-L2 task, the mean accuracy results are 0.376 for high variability condition, 0.338 for medium condition and 0.345 for no variability condition. On the other hand, in the L2-to-L1 recall task the results decrease from no to high variability condition, having the highest average the no variability condition (0.538) and the lowest the high condition (0.513).

The results on mixed models on accuracy based on variability for both tasks indicated that there was an effect of condition, revealing that acoustic variability improved learning but only in the high variability condition (High Variability: $\beta = 0.56$; SE = 0.27, $z = 2.10$, $p < 0.05$; Medium Variability: $\beta = -0.01$; SE = 0.27, $z = -0.05$, $p = 0.960$). There was a task effect ($\beta = 1.16$; SE = 0.12, $z = 9.68$, $p < 0.001$), showing higher performance in the translation than in the picture-to-L2.

In addition, an interaction between task and condition (CondH:task: β = -0.36; SE = 0.17, z = -2.14, $p < 0.05$) was obtained. We further followed this interaction by running a mixed model separated by task in order to compare the no variability condition with the variability ones—medium and high conditions—. In the picture-to-L2 task, the results revealed that no variability did not differ significantly from either medium variability (β = −0.04; SE = 0.13, z = −0.30, p = 0.764) or the high variability (β = 0.23; SE = 0.12, z = 1.87, p = 0.062), while the comparison between high and medium was significant (β = −0.27; SE = 0.13, z = −2.17, $p < 0.05$). These results are indicators of a significant effect due to the increase of variability from medium to high conditions. The medium variability condition does not seem to trigger any effect in the participants when it comes to learning pseudowords since the mean is very similar to the no variability condition.

On the other hand, results on accuracy based on variability for L2-to-L1 task indicated that the effect of variability was not significant. We did not observe any significant effect in the mixed model analysis comparing no variability and medium variability conditions (β = -0.07; SE = 0.12, z = -0.63, p = 0.530) nor comparing no variability and high variability conditions (β = -0.16; SE = 0.12, z = -1.40, p = 0.163).

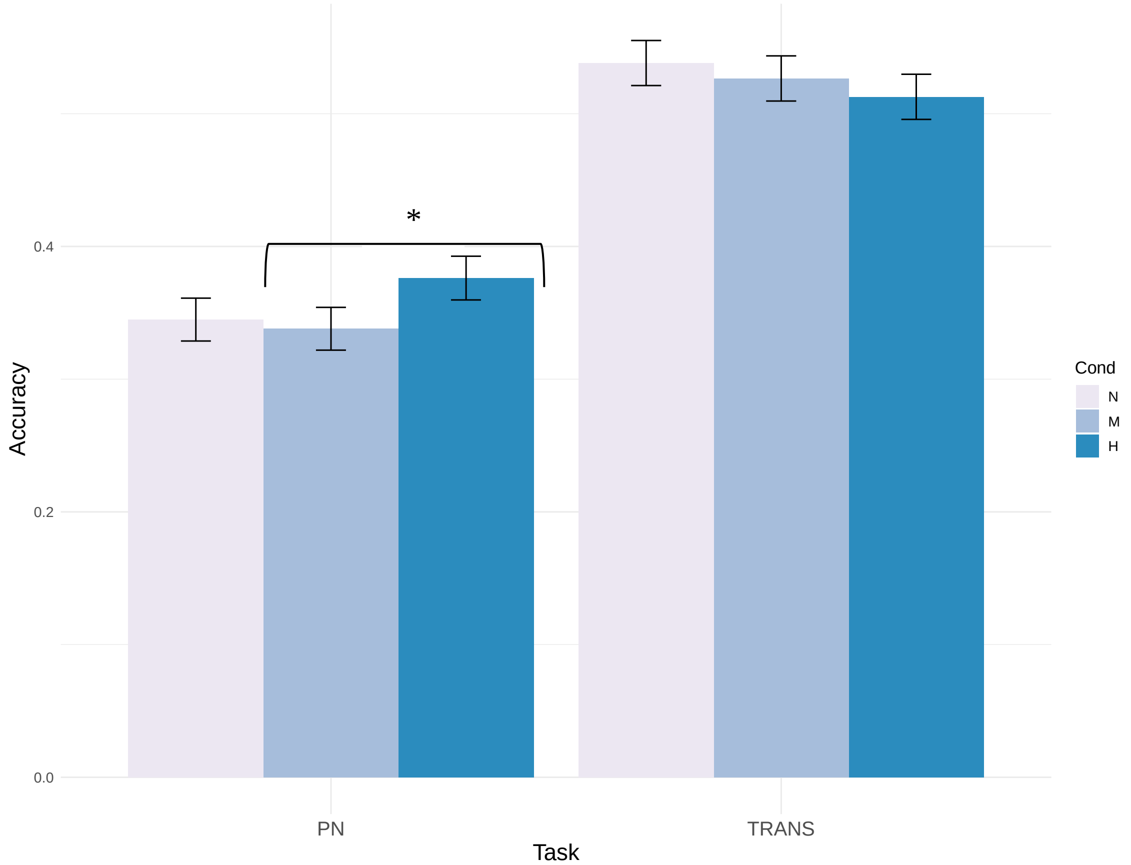

Figure 2. Effects of speaker variability on accuracy of picture-to-L2 and L2-to-L1 in experiment 1. Error bars represented. * $p < 0.05$.

In conclusion, a quite low accuracy was reported in the pseudowords learning tasks. Nonetheless, our results showed a tendency for improved learning with the increased variability, although this increase was more marked between the medium and the high variability conditions.

DISCUSSION

The findings from Experiment 1 can be summarized as follows. The high variability condition resulted in higher accuracy than the other variability conditions for picture-to-L2 task which shows a positive effect of acoustic variability on L2 vocabulary learning. As we obtained a significant effect on the comparison between high and medium conditions, our results support the idea that speaker variability can promote more robust lexical representations of new L2 vocabulary, hence reinforcing learning (Barcroft & Sommers 2005). However, they also suggest that there must be enough variability in the signal to get an improvement in learning. Furthermore, in the L2-to-L1 task we obtained the opposite pattern of results from what was expected, even though they are not significant, and the accuracy is higher than in the other task.

Overall, the general accuracy was quite low (relative to the original study) which can be due to pseudowords being more difficult to learn than words themselves or also because many more items were used in the present study (48 compared to 24). It is known that adult second-language learners' brain activity, as measured by event-related potentials (ERPs), discriminate between L2 words and L2 'pseudowords' (McLaughlin, Osterhout & Kim 2004). Previous studies have suggested that this increase is due to an increased demand on the lexical access system (Newman & Twieg 2001). These findings may potentiate what we have already hypothesised about pseudowords being more difficult to learn than words. Furthermore, it is known that learning to produce acoustically far L2 sounds through multiple speakers training is more effective than learning to produce close L2 sounds that might place major demands on learners' cognitive resources (Brosseau-Lapré et al. 2013; Kartushina & Frauenfelder 2014). The pseudowords that have been used in the experiment are not alike, which could have made it easier to learn. Despite that, pseudowords are not close to our first languages, Spanish and Catalan, which might be the reason for a low accuracy in the first place.

Finally, the significant effect between tasks could be explained as the L2-to-L1 task being much easier than the picture-to-L2 task. The reason for this could be that, in the L2-to-L1 task, the participants did not have to produce words which they were not familiar with (pseudowords); however, they did in the picture-to-L2 task.

Experiment 2

METHODS

Participants

In this experiment, a total of 54 subjects were recruited (14 males and 40 females), but we excluded 3 participants: one of them fell asleep during the learning task, the second one had visual difficulty with the original video and missed 5 of the words in the last variability condition, and the last one did not record himself during the learning exercise. For this reason, the analysis had ultimately had to be performed with just 51 participants (11 males and 40 females).

Materials

The initial stimuli consisted of 48 pictures, the ones used in the previous experiment, and their respective Catalan signs (LSC). The material and design (Appendix A) were the same as experiment 1. In this case, the words were also divided into three categories which contained 16 words each:

> **Word set 1**: *Molino, bombero, buzón, hucha, oveja, ajo, cebolla, calcetín, camello, cereza, hormiga, kiwi, lechuga, pepino, pera and limón.*
> **Word set 2**: *Araña, árbol, archivador, aspiradora, berenjena, bota, fresa, melón, guisante, hamburguesa, galleta, ciervo, langosta, tiburón, oliva and uva.*
> **Word set 3**: *Tigre, rana, patata, pastilla, muñeco, melocotón, abeja, espárrago, bombilla, piña, lagartija, monja, zorro, pimiento, pincel and regadera.*

Again, we used a soundproof room to gather the recordings of 7 different Catalan signers (3 males and 4 females). Six signers were used for the learning phase (3 males and 3 females), whereas the remaining one was used for the production tasks in order to ensure the use of novel signers. Each of them produced two lists of recordings so the researcher could choose the best input quality. Plus, each video was fragmented to a single mp4 file using Kdenlive, a free and open-source video editing software based on the MLT Framework, KDE and Qt.

Data analysis

In order to find a proper criterion to score accuracy in the picture-to-L2 recall, the three composing parameters of signs were considered: handshape, movement and location. These three parameters were selected because they are the most distinctive features and the ones that best determine how a sign is made (Sandler 2012).

Just if the participant performed three of them correctly, the equivalent accuracy was 1, if not 0; no responses were also labeled as 0. For the L2-to-L1 task, the same criterion was used. Finally, the results were analyzed with R studio and a mixed model analysis was run too (commands in Appendix B and C).

RESULTS

Firstly, means and SE for accuracy based on variability were calculated and depicted in Figure 3. As can be seen in the graph, the best accuracy corresponds to the medium variability condition (0.729) in the picture-to-L2 task. However, the accuracy is lower in the no variability condition (0.695) than in the higher one (0.717) in this task. On the other hand, in the L2-to-L1 exercise, the results range from lowest to highest through the conditions, having the highest average the high condition (0.913). All things considered, the accuracy is much better in this experiment than in the previous one, being the lowest accuracy (0.695) of this experiment 2 higher than the highest one (0.538) of experiment 1 and the highest of experiment 2 (0.913) quite close to 1.

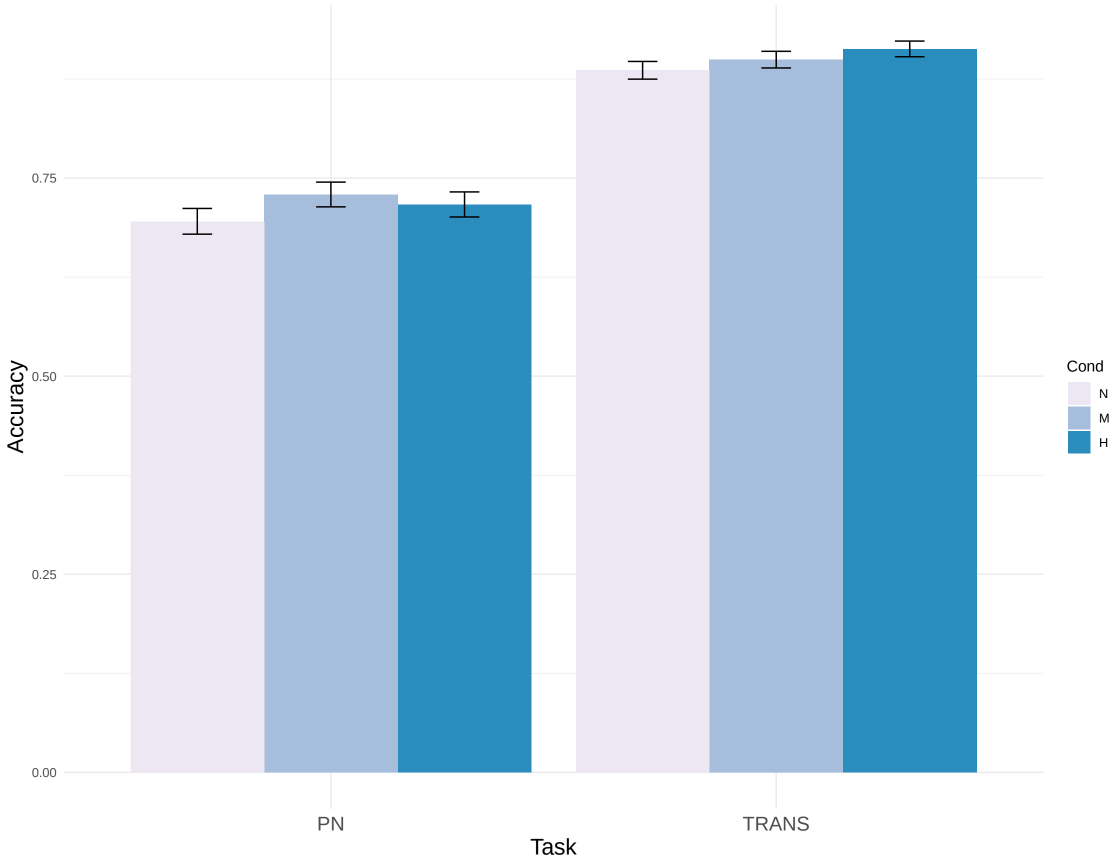

Figure 3. Effects of signer variability on accuracy of PN and L2-to-L1 in experiment 2. Error bars represented.

The mixed model's results on accuracy based on variability for both tasks did not show a condition effect, but they did show a task effect (β = 1.55; SE = 0.15, z = 10.32, $p < 0.001$). Moreover, the results indicate that there was not an interaction between task and condition (CondM:task: β = -0.02; SE = 0.22, z = -0.10, p = 0.917 and CondH:task: β = 0.23; SE = 0.22, z = 1.05, p = 0.296) showing that the number of signers did not affect accuracy. For this reason, we did not further study this interaction by running a mixed model separated by task.

In conclusion, no significant interaction between task and condition was found in experiment 2. Despite that, subjects were excellent sign learners. The reason could be due to a ceiling effect, which refers to the point at which an independent variable no longer has an effect on a dependent variable, once a kind of saturation has been reached; in this case, that would be the low difficulty of the experiment itself. For this reason, they were tested again two weeks later in order to check whether the effect of variability would be seen, which might indicate that variability affects short-term and long-term memory differently.

Retest

Firstly, means and SE for accuracy based on variability were calculated and depicted in Figure 4. As it can be seen in the bar plot, the results in the picture-to-L2 are 0.372 for no variability, 0.445 for medium variability and 0.424 for high variability. In what regards to the L2-to-L1 recall task, the results are 0.792 for no variability, 0.816 for medium variability and 0.796 for high variability. To summarize, the accuracy for the L2-to-L1 task is similar to the results in the first sign experiment, but it is lower in the picture-to-L2 task.

The mixed model's results on accuracy based on variability for both tasks indicated that there was an effect on condition (Medium Variability: β = 0.62; SE = 0.32, z = 1.94, p = 0.05 and High Variability: β = 0.62; SE = 0.32, z = 1.93, p = 0.05) and a task effect (β = 2.46; SE = 0.16, z = 15.55, $p < 0.001$). The interaction between task and condition was not significant (CondM:task: β = -0.26; SE = 0.21, z = -1.20, p = 0.229 and CondH:task: β = -0.34; SE = 0.21, z = -1.60, p = 0.110), revealing similar results in both tasks.

In conclusion, a significant effect was observed in the picture-to-L2 task for accuracy between no variability and variability conditions. This significant effect is due to the increase of variability. Moreover, similar accuracy was found in the L2-to-L1 task compared to that in experiment 2, but lower accuracy results were found in the picture-to-L2 task within the retest, that being compared to the original experiment 2.

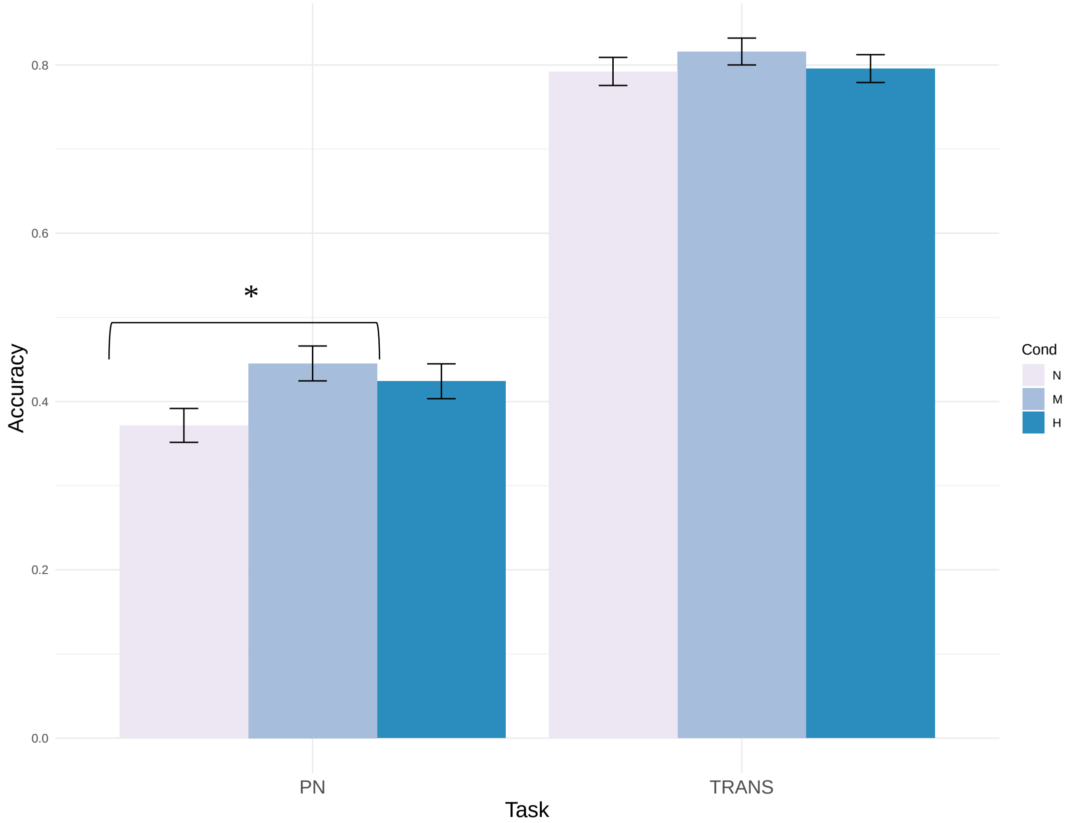

Figure 4. Effects of signer variability on accuracy of PN and L2-to-L1 in experiment 2 retest. Error bars represented. * $p < 0.05$.

DISCUSSION

The findings in experiment 2 can be summarized as follows. There is no positive effect of variability on short-term LSC learning, so our initial hypothesis was not full supported. Nonetheless, the results of the retest showed that participants remember fewer signs after two weeks, but their long-term LSC learning is affected by variability. We can state that because results in the picture-to-L2 task were better for high and medium conditions, and a significant effect was found between no and medium conditions. Taking these findings into account, we conclude that variability does improve the learner's sign production.

Previous studies show that variability improves oral second language learning and motor learning, but there are no studies about variability in sign languages or in visual data. Despite that, our initial hypothesis remains consistent with the motor learning improvement due to variability since there is an effect of variability in the picture-to-L2 task which is the one in which they have to produce the sign, but only in the retest experiment. This finding suggest that variability helps when the learning context is difficult as we know that the general accuracy is lower in the retest (Dhawale et al. 2017).

GENERAL DISCUSSION

The objective of the present thesis was to explore how acoustic variability, indexed by the number of teachers, influences second language learning and how learning is modulated by the similarity of the language to be learned with the first language of the learner.

All things considered, the first result we observed was that accuracy in LSC experiments is much better than in the pseudowords experiment, and this can be due to different factors. A possible explanation for this could be that some iconic or transparent relations strengthen lexico-semantic connections. For example, if the sign reminds the learner the properties of an object, the access to semantics is more reinforced than if he has to learn an invented word that is not similar to the same word in his language. In this case, it could be possible that learners were not influenced by variability since they may have another more robust way to hold on. Moreover, an enactment effect could be the reason why signs are easier to learn than pseudowords. This effect describes the fact that learners can memorize better if they perform the described action during learning, which refers to learning by doing (Russ et al. 2003). This may indicate that the effectiveness of variability training depends, not only on individuals' aptitude, but also on the nature of the categories being acquired (Sadakata & McQueen 2014). Despite that, we cannot ensure that this happened in our experiments because we did not look at the whole learning phase of the experiments since this was not our focus of interest.

In addition, it is known that sign recognition appears to follow a systematic pattern in which information about the location of the sign is reliably identified at first, then followed by handshape information and finally the movement. Besides, despite the slower articulation of a sign when compared to a word, sign recognition appears to be faster than word recognition. It has been observed that proportionally less of the signal needs to be processed in order to uniquely identify a sign compared to a spoken word (Corina 2015).

To sum up, regarding our first two questions, variability improves L2 learning in the picture-to-L2 task. In our first experiment, the high variability condition produced a positive effect. Moreover, the short-term LSC learning is not affected by variability, but the long-term is. This could prove that variability enhances motor learning. On the other hand, LSC is easier to learn than pseudowords due to the similarity with our first language (semantic relations) or due to an enactment effect. In both experiment 2 and retest we found a higher accuracy than in experiment 1.

These results show that variability enhances second language learning in different modalities regardless of the similarity of the language to be learned with the first language of the learner. Variability only affects picture-to-L2 task which may be because in this task the pictures used are the same as in the learning phase while in the L2-to-L1 recall the audios used are different from the audios of the learning phase. Even though, we do not know why there is an effect in the Barcroft and Sommers study in this task but there is not in our experiments.

BIBLIOGRAPHY

Barcroft, J., 2001, “Acoustic variation and lexical acquisition,” *Language Learning*, 51(4), 563–590.

Barcroft, J. & Sommers, M.S., 2005, “Effects of acoustic variability on second language vocabulary learning,” *Studies in Second Language Acquisition*, 27(3), 387–414.

Brosseau-Lapré, F., Rvachew, S., Clayards, M. & Dickson, D., 2013, “Stimulus variability and perceptual learning of nonnative vowel categories,” *Applied Psycholinguistics*, 34(3), 419–441.

Corina, D.P., 2015, *Sign Language: Psychological and Neural Aspects*, Second Edi, vol. 21, Elsevier.

Dhawale, A.K., Smith, M.A. & Ölveczky, B.P., 2017, “The Role of Variability in Motor Learning,” *Annual Review of Neuroscience*, 40(1), 479–498.

Dong, H., Clayards, M., Brown, H. & Wonnacott, E., 2019, “The effects of high versus low talker variability and individual aptitude on phonetic training of Mandarin lexical tones,” *PeerJ*, 2019(8), 1–45.

Giannakopoulou, A., Brown, H., Clayards, M. & Wonnacott, E., 2017, “High or low? Comparing high and low-variability phonetic training in adult and child second language learners,” *PeerJ*, 2017(5), 1–35.

Goldinger, S.D., Pisoni, D.B. & Logan, J.S., 1991, “On the Nature of Talker Variability Effects on Recall of Spoken Word Lists,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(1), 152–162.

Hardison, D.M., 2003, “Acquisition of second-language speech: Effects of visual cues, context, and talker variability,” *Applied Psycholinguistics*, 24(4), 495–522.

Jaeger, T.F., 2008, “Categorical data analysis: Away from ANOVAs,” *Journal of Memory and Language*, 59(4), 434–446.

Kartushina, N. & Frauenfelder, U.H., 2014, “On the effects of L2 perception and of individual differences in L1 production on L2 pronunciation,” *Frontiers in Psychology*, 5(NOV), 1–17.

Lively, S.E., Logan, J.S. & Pisoni, D.B., 1993, “Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories,” *The Journal of the Acoustical Society of America*, 94(3), 1242–1255.

McLaughlin, J., Osterhout, L. & Kim, A., 2004, “Neural correlates of second-language word learning: Minimal instruction produces rapid change,” *Nature Neuroscience*, 7(7), 703–704.

Newman, S. & Twieg, D., 2001, “Differences in Auditory Processing of Words and Pseudowords: An fMRI Study,” *Human Brain Mapping*, 14, 39–47.

Russ, M.O., Mack, W., Grama, C.R., Lanfermann, H. & Knopf, M., 2003, “Enactment effect in memory: Evidence concerning the function of the supramarginal gyrus,” *Experimental Brain Research*, 149(4), 497–504.

Sadakata, M. & McQueen, J.M., 2014, “Individual aptitude in Mandarin lexical tone perception predicts effectiveness of high-variability training,” *Frontiers in Psychology*, 5(NOV), 1–15.

Sandler, W., 2012, “The Phonological Organization of Sign Languages,” *Linguistics and Language Compass*, 6(3), 162–182.

Sinkeviciute, R., Brown, H., Brekelmans, G. & Wonnacott, E., 2019, “THE ROLE of INPUT VARIABILITY and LEARNER AGE in SECOND LANGUAGE VOCABULARY LEARNING,” *Studies in Second Language Acquisition*, 41(4), 795–820.

Tagarelli, K.M., Ruiz, S., Vega, J.L.M. & Rebuschat, P., 2016, “Variability in second language learning,” *Studies in Second Language Acquisition*, 38(2), 293–316.

Appendix A

Group	VARIABILITY	SPEAKERS	SET 1	SET 2	SET 3
	N	1	1	2	3
1	M	1,2,4	2	3	1
	H	1,2,4,3,5,6	3	1	2
	N	2	1	2	3
2	M	2,3,5	2	3	1
	H	2,3,5,1,4,6	3	1	2
	N	3	1	2	3
3	M	3,1,6	2	3	1
	H	3,1,6,2,4,5	3	1	2
	N	4	1	2	3
4	M	4,5,1	2	3	1
	H	4,5,1,2,3,6	3	1	2
	N	5	1	2	3
5	M	5,6,2	2	3	1
	H	5,6,2,1,3,4	3	1	2
	N	6	1	2	3
6	M	6,4,3	2	3	1
	H	6,4,3,1,2,5	3	1	2

Appendix B

```
setwd("~/Desktop/TFG/LSC/R")
library(readxl)
dataR <- read_excel("expe1-R.xlsx", sheet = "R-EXP1")
summary (dataR)
names (dataR)
str(dataR)
dataR$pict = as.factor(dataR$pict)
dataR$nonword = as.factor(dataR$nonword)
dataR$task = as.factor(dataR$task)
dataR$cond = as.factor(dataR$cond)
dataR$order = as.factor(dataR$order)
dataR$subj = as.factor(dataR$subj)
dataR$List = as.factor(dataR$List)
My_Theme = theme(axis.title.x = element_text(size = 16),axis.text.x = element_text(size
= 14),axis.title.y = element_text(size = 16))
library(ggplot2)
library(plyr)
MeanGRdataR <- ddply(dataR, c("task", "cond"), summarise,
        N    = length(accuracy),
        mean = mean(accuracy),
        sd   = sd(accuracy),
        se   = sd / sqrt(N))
Library (dplyr)
BarR<- MeanGRdataR %>%
  mutate(cond = factor(cond, levels=c("N", "M", "H"))) %>%
  group_by(task,cond) %>%
  ggplot(aes(x = task, y = mean, fill = cond)) +
  geom_bar(stat="identity", position=position_dodge())+
  geom_errorbar(aes(ymin=mean-se, ymax=mean+se), width=.2,
        position=position_dodge(.9))+
  labs(x="Task", y="Accuracy")+
  guides(fill=guide_legend(title="Cond"))+
  scale_fill_brewer(palette="PuBu")+
  theme_minimal()+
  My_Theme
print (BarR)
```

Appendix C

```
setwd("~/Desktop/TFG/LSC/R")
library(readxl)
library(magrittr)
library(dplyr)
library(lme4)
dataR <- read_excel("expe1-R.xlsx", sheet = "R-EXP1")
summary (dataR)
names (dataR)
str(dataR)
dataR$pict = as.factor(dataR$pict)
dataR$nonword = as.factor(dataR$nonword)
dataR$cond = as.factor(dataR$cond)
dataR$order = as.factor(dataR$order)
dataR$subj = as.factor(dataR$subj)
dataR$List = as.factor(dataR$List)
dataR$task<-as.numeric (dataR$task)
 ###Linear mixed model########
lmerVarsuj <- glmer(accuracy ~ cond*task + (1 | subj) + (1 | pict), data= dataR, family=binomial)
summary(lmerVarsuj)
PN<-subset (dataR, task=="1")
lmerVarPN <- glmer(accuracy ~ cond + (1 | subj) + (1 | pict), data= PN, family=binomial)
summary(lmerVarPN)
TRANS<-subset (dataR, task=="2")
lmerVarTRANS <- glmer(accuracy ~ cond + (1 | subj) + (1 | pict), data= TRANS, family=binomial)
summary(lmerVarTRANS)
```