ASM Final Assignment: Phonotactic similarity Study

Zulikah Latief S4273281
Date: 3rd February, 2022

Methods

Research Aim

The study focuses on two groups of subjects, native Dutch and native Polish, who were trained to learn an artificial language. The phonotactic rules of the language were more similar to Polish than Dutch. In other words, the words of the artificial language looked more like Polish words than Dutch words. The aim of the research is to investigate the influence of phonotactic similarity with the native language when learning a new language; i.e., does phonotactic similarity help or hinders learning a new language.

Participants

The study consists of 60 participants, 30 native Dutch speakers and the remaining 30 native Polish speakers. Note that neither of the subject groups were told about the phonotactic similarity of the artificial language with Polish.

Procedure and Experimental Design

The experiment is divided into two stages, namely, *pre* test stage and a *post* test stage wherein subjects are given artificially generated letter strings (items) of *Type* word and non-word. Note that there is also an additional training stage between the pre test and post test wherein each subject is given a list of words from the artificial language for practice. However, no data about this session is made available. There are a total of 120 *Items*, 60 items in each *Test* stage. Note that the items across the two Test stages are different and not all items have equal number of repeated measures. Each subject's accuracy was recorded for correctly identifying the *Type* (word/ non-word) of the presented item. During each of the Test stage (pre and post), each participant undergoes 30 trials (15 words and 15 non-words) and their accuracy is recorded. Thus, each subject has 60 trials (30 pretest and 30 post test), which makes a total of, 60 trials * 60 subjects = 3600 trials in the whole experiment. However, subjects s048 and s049 (Polish and Dutch respectively) were not a part of the pretest stage, and thus each undergo a total of 30 trials in the post-test only. Therefore, there are a total of 3540 trials (60 trails for each of 58 subjects and 30 trials for each of the 2 subjects). Half of these trials, i.e., 1770 were of type word and the rest were of the type non-word.

The design of the experiment in terms of within-subject and between-subject is not fixed. For instance, with respect to the variable Type, each subject is presented with both levels of the variable (word and nonword) which makes it within-subject. However, not all subjects take part in both pre and post test conditions (s048 and s049 are excluded from pre-test). Moreover, with respect to the variable *Language*, each subject can only be in one of the levels (Dutch or Polish), which makes it a between-subject study.

Analysis GLME

Analysis technique

The response variable in the study is *Accuracy*, which is a binary variable. This type of data is always modelled using logistic regression. Therefore, a generalised linear model is appropriate. Moreover, *Subjects* and *Items* are repeated measures that can be generalised over. A mixed-effect model such as the GLME can handle repeated measures and binomial data, therefore a GLME is used.

Description of predictors

The experimental manipulations such as Type (word and nonword), Test (pre and post) and Language (Dutch and Polish) are used as fixed effects for the GLME analysis. Random intercepts for subjects and items are chosen since they have numerous repeated measures. Random slope for Type is considered subjects may vary in their accuracy across different levels of type (word and nonword). Similarly, since subjects may vary in their accuracy across different Test, i.e., pre and post test, random slope for Test was also considered.

Description of analysis procedure.

Firstly, a maximal fixed effect model is considered as the base model. Through forward fitting random effects (intercepts followed by slopes) are added. Model comparison, visualizing residuals and checking for correlation are three steps used in determining the final random effect structure. Similarly, fixed effects structure is determined through backward fitting that starts from a maximal model. Similar to random effects, fixed effect structure is determined through model comparison, summary inspection and visualization. Finally, the best fitting model is evaluated by checking residuals and extracting model estimates.

Results

Data

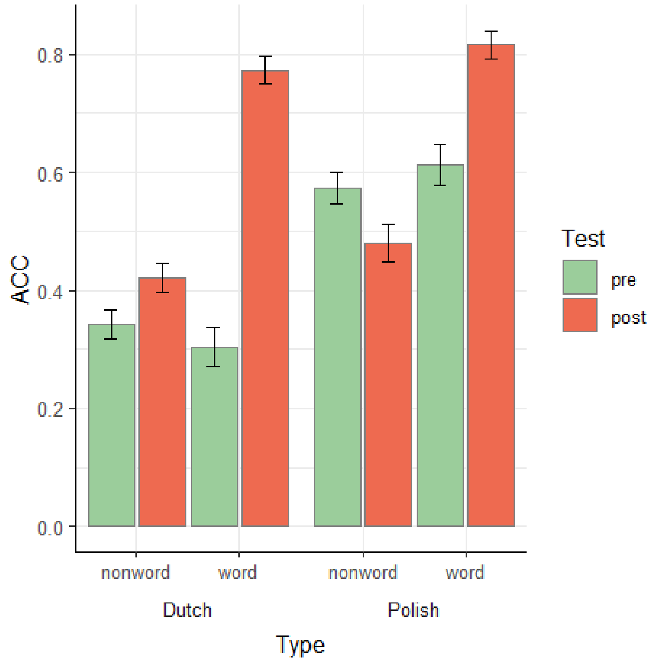

Figure 1: Mean Accuracy of Subjects across Test, Type and Language

Figure 1 provides a summary of the data, The bar plot shows the mean accuracy of Dutch and Polish subjects in the lexical decision task. For native Dutch speakers, the mean accuracy for type word and non-word during the pre test stage is lower when compared to post test stage. The mean accuracy for non-word type increases by a relatively small factor of 0.1 in the post test stage. Interestingly, for type 'word', there is a dramatic increase in mean accuracy from 0.3 (pre-test) to 0.75 (post-test). Overall, the increase in mean accuracy from pre to post test is visible in both 'word' and 'non-word' type, however, the increase is highly significant for type 'word'.

For native Polish speakers, the mean accuracy of type non-word is high at 0.57 during the pre-test stage. However, this accuracy falls to 0.49 during the post-test stage. For type 'word' there is a significant increase of 0.2 in mean accuracy from pre test to post test stage. However, this increase is not as dramatic as seen in Dutch native speakers.

Overall, native Polish speakers show a higher mean accuracy for all 'Type' and 'Test' categories. Specifically, native Polish speakers have a significantly higher mean accuracy for both 'word' and 'non-word' in the pre-test when compared to native Dutch speakers. However, the increase in mean accuracy during the post test stage for both categories of 'Type' is more significant for Dutch speakers. Moreover, the accuracy drops rather than increases for native Polish speakers during the post-test stage for type 'non-word'.

The data set contained 106 NaN values which were removed as they accounted for a small percentage of the total trials.

Analysis of result The best fitting model contains $ACC\ Language + Type + Test + Type : Test + Language : Test + (1 + Test|Subject) + (1|Item)$. The fitted vs residual plot (refer figure 2) shows a clear pattern wherein extreme accuracy values have higher number of residuals that are low in magnitude. As the accuracy goes from 0 to 0.5, the residuals go from highly concentrated small values to sparse and large negative values. At 0.6 accuracy, there is a flip in the residuals. The residuals are high in magnitude and positive and become more concentrated and low in magnitude towards accuracy value 1. There is a clear distinction of two trends. Some of these trends in the residuals can be seen in the residuals of the fixed effects. Therefore, I suspect there is either correlation between the variables or there is a missing fixed effect.

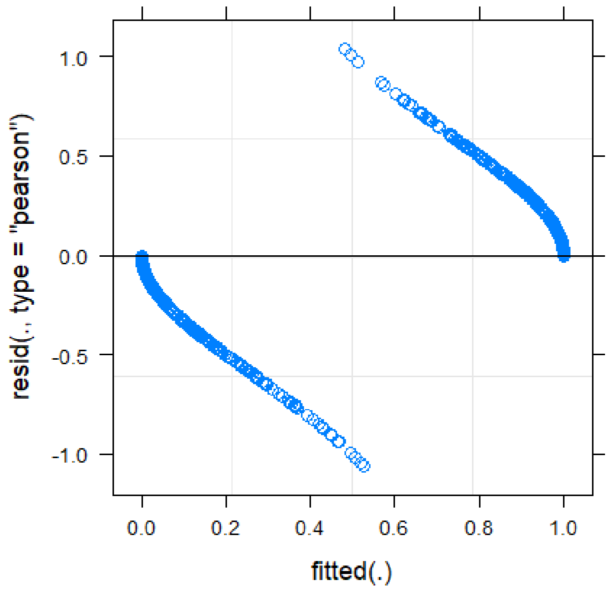

Figure 2: Residuals vs fitted for best fitting model GLME

Finally, model estimates from the best fitting model for each condition is extracted and plotted as shown in figure 3. In the figure, Dutch native speaker has accuracy of 0 for non-words across both pre and post test stage. However, there is a slight increase in accuracy during the post test stage. Whereas, for native polish speaker, pre-test stage has a high accuracy of 1 for nonwords and this drops drastically during the post test stage. For the word type, native Dutch speaker starts of with an accuracy of 0 during the pre test but after the training the accuracy shoots to 1 during the post test. On the other hand, for the native polish speaker the accuracy for word type remains consistent across both pre and post test stage.

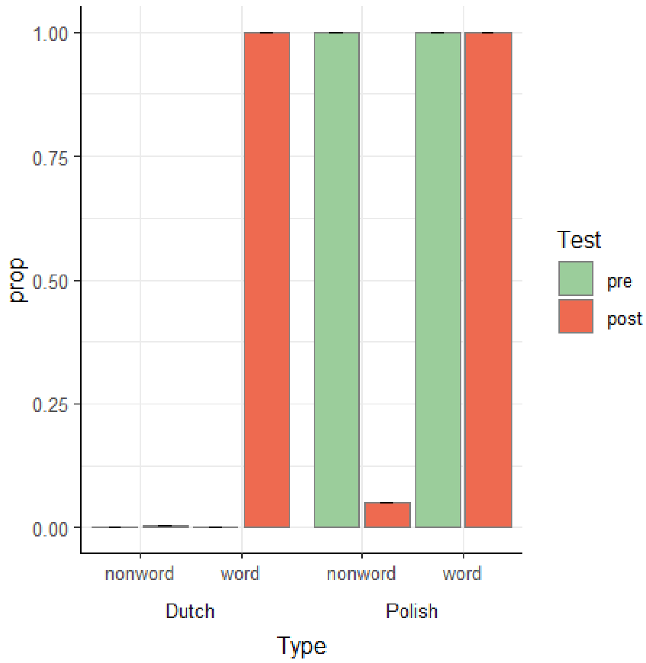

Figure 3: Model estimates retrieved for all conditions

Based on the thorough analysis, my conclusion is that although Polish native speakers have an overall higher accuracy, it is more difficult for them to learn the artificial language. We know that the words presented are more similar to Polish. During the pre test stage the accuracy for both words and nonwords are high for native polish speakers. However, after the training process, the nonword accuracy drops drastically. This maybe because the nonwords are probably not similar to Polish (unlike the words) and this might confuse the subjects. It could also be that nonwords in general result in lower accuracy. However, from the data (figure 1) we know that the accuracy of nonwords for native Dutch increases after the training. This is not the case for native Polish speakers. The underlying research focuses on understanding whether or not Polish speakers have an advantage in learning the artificial language. Since, their accuracy drops for nonwords (rather than increase), it suggests that the Polish native speakers do not have an advantage. They infact might be at the disadvantage. On the other hand, Dutch speakers start out with low accuracy for both words and non-words but after training their accuracy either increases or remains the same, but it does not decrease.