Universiteit Utrecht

Faculty of Humanities
Version September 2014

PLAGIARISM RULES AWARENESS STATEMENT

Fraud and Plagiarism
Scientific integrity is the foundation of academic life. Utrecht University considers any form of scientific deception to be an extremely serious infraction. Utrecht University therefore expects every student to be aware of, and to abide by, the norms and values regarding scientific integrity.

The most important forms of deception that affect this integrity are fraud and plagiarism. Plagiarism is the copying of another person's work without proper acknowledgement, and it is a form of fraud. The following is a detailed explanation of what is considered to be fraud and plagiarism, with a few concrete examples. Please note that this is not a comprehensive list!

If fraud or plagiarism is detected, the study programme's Examination Committee may decide to impose sanctions. The most serious sanction that the committee can impose is to submit a request to the Executive Board of the University to expel the student from the study programme.

Plagiarism
Plagiarism is the copying of another person's documents, ideas or lines of thought and presenting it as one's own work. You must always accurately indicate from whom you obtained ideas and insights, and you must constantly be aware of the difference between citing, paraphrasing and plagiarising. Students and staff must be very careful in citing sources; this concerns not only printed sources, but also information obtained from the Internet.

The following issues will always be considered to be plagiarism:

- cutting and pasting text from digital sources, such as an encyclopaedia or digital periodicals, without quotation marks and footnotes;
- cutting and pasting text from the Internet without quotation marks and footnotes;
- copying printed materials, such as books, magazines or encyclopaedias, without quotation marks or footnotes;
- including a translation of one of the sources named above without quotation marks or footnotes;
- paraphrasing (parts of) the texts listed above without proper references: paraphrasing must be marked as such, by expressly mentioning the original author in the text or in a footnote, so that you do not give the impression that it is your own idea;
- copying sound, video or test materials from others without references, and presenting it as one's own work;
- submitting work done previously by the student without reference to the original paper, and presenting it as original work done in the context of the course, without the express permission of the course lecturer;
- copying the work of another student and presenting it as one's own work. If this is done with the consent of the other student, then he or she is also complicit in the plagiarism;
- when one of the authors of a group paper commits plagiarism, then the other co-authors are also complicit in plagiarism if they could or should have known that the person was committing plagiarism;
- submitting papers acquired from a commercial institution, such as an Internet site with summaries or papers, that were written by another person, whether or not that other person received payment for the work.

The rules for plagiarism also apply to rough drafts of papers or (parts of) theses sent to a lecturer for feedback, to the extent that submitting rough drafts for feedback is mentioned in the course handbook or the thesis regulations.
The Education and Examination Regulations (Article 5.15) describe the formal procedure in case of suspicion of fraud and/or plagiarism, and the sanctions that can be imposed.

Ignorance of these rules is not an excuse. Each individual is responsible for their own behaviour. Utrecht University assumes that each student or staff member knows what fraud and plagiarism

Universiteit Utrecht

entail. For its part, Utrecht University works to ensure that students are informed of the principles of scientific practice, which are taught as early as possible in the curriculum, and that students are informed of the institution's criteria for fraud and plagiarism, so that every student knows which norms they must abide by.

I hereby declare that I have read and understood the above.
Name: Christiaan Alex Brouwer Student number: 3829820
Date and signature: 23/7/2017 

Submit this form to your supervisor when you begin writing your Bachelor's final paper or your Master's thesis.

Failure to submit or sign this form does not mean that no sanctions can be imposed if it appears that plagiarism has been committed in the paper.

The effect of music training on the processing of the prosody-semantics interface by Dutch learners of English

BA Thesis

Alex Brouwer, 3829820

Supervisor: Aoju Chen

Second reader: Iris Mulders

Geuzenpad 14, 2722AV Zoetermeer

23 July 2017

6158 words

Abstract

The interface between prosody and semantics is utilised to determine semantic meaning using prosodic information. Recently, more research has been devoted to the workings of this interface. However, not much research has gone into this interface in second language (L2) learners. One variable that could potentially help L2 learners of English is music training. Music training has been associated with benefits to prosodic perception and processing in the literature, also in foreign languages. In the present study, we have looked at how Dutch speakers of L2 English with varying amounts of music training and listening use English prosodic information to understand semantic meaning of utterances with 'only'. We tested 30 participants on their online processing of prosodic information in utterances with 'only' in a visual world paradigm experiment using eye-tracking. Generally, Dutch learners of English do not use prosodic information as effectively as predicted. More importantly, we have found that music training had no effect on the perception and processing of prosodic information. Confirmation of the null-hypothesis may, in this case, have been the result of some methodological issues that could and should be considered for future research. However, our results in combination with recent literature might also suggest that music training may not actually have the benefits in this specific usage of prosody.

Table of Contents

Acknowledgments 4

1. Introduction 5

    1.1 ‘Only’ in Focus 5

    1.2 Music Training and Prosody in L2 8

    1.3 Listening to Music 9

    1.4 Current Study and Hypothesis 9

2. Method 11

    2.1 Participants 11

    2.2 Materials 12

    2.3 Procedure 14

3. Predictions 15

4. Results 16

    4.1 Preliminary Remarks 16

    4.2 Multiple Linear regression 19

5. Discussion 20

6. Conclusions and Future Research 24

References 25

Appendix A 28

Appendix B 30

Acknowledgements

Firstly, I would like to thank my supervisor Aoju Chen for all her help with this thesis and for introducing me to the field of prosody in the first place.

Secondly, thanks to Emily Ge for allowing me to be a part of such interesting research, helping me with my thesis and teaching me many new skills.

I would also like to thank Rachida Ganga for the help she provided as well as for the occasional emotional and moral support.

Lastly, I would like to thank my fiancé Caitlin for pretending to listen to my ramblings and her love and support.

1. Introduction

1.1 'Only' in Focus

In West-Germanic languages such as Dutch and English, the focus of an utterance can be marked prosodically by use of pitch accent. Using pitch accents in such a way, the speaker can convey to the listener what the prominent part of the utterance is. This means that prosodic information can be utilised by speakers to help listeners understand the information status of words and utterances. In this way, pitch accents can identify new or contrastive information. The example in (1b), for instance, shows how a pitch accent can be used to contrast with previously held assumptions (capitalisation marks accented words). In the case of (1b), the pitch accent serves a pragmatic purpose rather than a semantic one, because neither the meaning nor the interpretation, of the utterance in (1b) changes without a pitch accent on 'carried'. The only purpose the pitch accent on 'carried' has in (1b) is pragmatic because it is used to accentuate the contrast between 'kicking' and 'carrying' in the context presented by the question in (1a).

1. a) Did the dinosaur kick the ball?

   b) No, the dinosaur CARRIED the ball.

The example in (1) shows that prosodic information can be used pragmatically. However, there are also instances where it contributes to the semantics of an utterance. Rooth (1992) explored the semantics of sentences with the focus particle 'only'. He explained that in sentences with 'only' as a focusing adverb the truth condition of a sentence depends on which constituent receives focus. For instance, in a scenario where the dinosaur kicked both the suitcase and the ball to the cat, the sentence in (2a) with 'only' and focus marked by a pitch accent on 'cat' would be perfectly felicitous. However, the example in (2b) would not be

felicitous in the given context because of the focus particle ‘only’ and focus marked with a pitch accent on ‘ball’. The fact that the sentence in (2b) is infelicitous is because ‘only’ adds a focal meaning component to the sentence. In (2a), the focal meaning component is: ‘the dinosaur kicked the ball to no one other than the cat’. This means that the dinosaur could still also have kicked the suitcase to the cat, since the use of ‘only’ in combination with a pitch accent on ‘cat’ does not exclude that as a possibility. This makes the example in (2a) perhaps a bit of an odd, but true description of the situation. In the case of (2b), the focal meaning component is: ‘the dinosaur kicked nothing other than the ball to the cat’. Since the aforementioned suitcase in the scenario is not a ball, the sentence in (2b) is no longer a felicitous description because of its focal meaning component. As a focusing particle, ‘only’ is dependent on its corresponding focus bearing element for its interpretation (Rooth, 1992).

2. a) The dinosaur only kicked the ball to the CAT.

b) #The dinosaur only kicked the BALL to the cat.

Because of this dependency, there is some ambiguity in sentences with ‘only’. At the moment speakers hear the word ‘only’, they are primed to hear a focus bearing element which needs to be resolved (Dimitrova, Stowe, Redeker, & Hoeks, 2010). When the focus bearing element is produced, the ambiguity is resolved and the sentence can be interpreted as it was intended (Gualmini, Maciukaite, & Crain, 2003). For instance, in the example in (3), it is not clear what the dinosaur is doing and not doing until the listener hears what part of the sentence receives focus by way of an emphatic pitch accent. In speech, focus can be marked by way of pitch accents. This means that in such cases the full semantic meaning of the sentence is partially determined by prosodic information. The prosodic prominence brought by the presence of a pitch accent is necessary to determine the semantic meaning and resolve

the ambiguity (Gualmini et al., 2003). Moreover, it has been found in Dutch speakers that such prosodic information is processed instantly (Mulders & Szendrői, 2016).

3. a) The dinosaur had a ball and a suitcase. He was going to kick and carry them. Then he changed his mind.
   b) The dinosaur is only carrying the suitcase.
   c) The dinosaur is only CARRYING the suitcase (not kicking the suitcase).

In regard to the workings of focus and 'only', Dutch and English are fairly similar (Bouma, Hendriks, & Hoeksema, 2007). However, there are some differences which could affect the interpretation of ambiguity somewhat. In English, it is possible to place 'only' before or after a verb when the focus bearing element is a noun following the verb as in the examples in (4). Syntactically, 'only' c-commands the focus bearing element in both cases, which is a necessary condition for focusing adverbs (Bouma et al., 2007). This means that 'only' cannot be placed post-verbally when the focus bearing element is the verb because then the condition is not met. It seems that ambiguity could be avoided easily by placing 'only' directly before the noun if this is the focus bearing element, as in (4b). However, English speakers prefer the pre-verbal position for 'only' when the noun receives focus (Bouma et al., 2007). In contrast, in Dutch, it is much more common for 'only' to directly precede the focus bearing element (Foolen, Van Gerrevink, Hogeweg, & Prawiro-Atmodjo, 2009). If Dutch learners of English use some language transfer when learning and using English, this could perhaps make it a more difficult to use prosodic information in English effectively since, in Dutch, they can expect the focused constituent to follow 'only' immediately more frequently than in English.

4. a) The dinosaur only kicks the BALL.

   b) The dinosaur kicks only the BALL.

1.2 Music Training and Prosody in L2

To perceive and process prosodic information in speech, language users need to pick up on its prosodic cues. Previous literature suggests that music training might be beneficial when it comes to processing prosodic cues. For instance, Schön, Magne, and Besson (2004) found that musicians utilise their superior musical abilities to process prosodic information. They showed that musicians more accurately noticed pitch violations in music and, more importantly, in speech. Similarly, Moreno et al. (2009) showed that 6 months of musical training improved the ability in 8-year-old children to detect small pitch variations in speech. Furthermore, musicians have also been found to interpret prosodic information more accurately than non-musicians in experiments on the interpretation of emotional state based on speech prosody (Thompson, Schellenberg, & Ilie, 2004). Lastly, Schellenberg and Moreno (2009) found that musicians are better at perceiving differences in pitch in short tones which suggests that musicians process the difference in pitch faster than non-musicians.

Besides advantages of music training on the perception of prosodic cues in a first language, some research has indicated that music training can also benefit speech perception in less familiar languages. Swaminathan and Gopinath (2013) found that music training benefits Indian children in the comprehension and vocabulary of L2 English. More specifically regarding prosody, Thompson et al. (2004) asked participants to judge emotional state of speakers based on utterances in a language that was unfamiliar to them. Again, musically trained participants outperformed the participants that were not musically trained. Moreover, Marques, Moreno, Castro, and Besson (2007) showed that musicians are also better at detecting variations in pitch in a foreign language. In their ERP experiment, they

found that French musicians were better at detecting small pitch differences in Portuguese than non-musicians.

1.3 Listening to Music

Some have suggested that even without music training, listening to music can be used to train certain musical abilities such as perceiving and understanding structures in music compositions (Bigand & Pouline-Charronnat, 2006). Moreover, Müllensiefen, Gingras, Musil, and Stewart (2014) found that actively engaging with music can have a positive effect on music skills such as beat perception. They suggest that listening to music actively could lead to enhancing music abilities. Honing and Lading (2009) also found that exposure to music can help participants to make judgements on music compositions and suggested that music listeners have many of the capabilities that are often associated with musicians. However, others have argued that the effects of music listening only last for a short amount of time and have no long lasting benefits (Rauscher & Hinton, 2006). This is generally called the *Mozart effect* and it is mostly associated with advantages in visuospatial abilities. It has been suggested that the only advantage music listening has on cognitive abilities is because music improves the mood of the listener. In turn, this would help to score better on cognitive tests (Schellenberg & Weiss, 2013).

1.4 Current Study and Hypothesis

The goal of the current study is to explore whether music training and actively listening to music gives Dutch speakers an advantage when integrating prosodic information (pitch accents) to understand the semantics of our experimental utterances in English. As mentioned before, sentences with ‘only’ as a focus particle carry two components with them: a non-focal meaning component and a focal meaning component. In the example in (5a), the focus lies with the verb ‘kicking’. The sentences in (5b) and (5c) are the non-focal meaning component and focal meaning component, respectively. In sentences like the one in (5a),

'only' as a focus particle creates certain expectations, as well as some ambiguity since the focus could also lie with 'ball'. Upon hearing 'only', it would be ambiguous which constituent will receive focus.

5. a) The dinosaur is only KICKING the ball.

   b) Non-focal: The dinosaur is kicking the ball.

   c) Focal: The dinosaur is not doing x (if x =/ kicking) to the ball

In our experiment, we wanted to see whether musically trained participants are better at perceiving and processing prosodic cues (pitch accents) to understand the non-focal and focal meaning component, and to resolve the ambiguity. To this end, we presented them with pictures depicting the non-focal meaning component and two pictures depicting a contrast to the two focal meaning components for the sentence with a pitch accent on the verb (= the dinosaur not KICKING the ball, but doing something else to it) and for the sentence with a pitch accent on the object (= the dinosaur not kicking the BALL, but kicking something else). Once the participants heard the necessary prosodic information, they could understand the focal meaning component. We could then see whether the participants increasingly looked at the pictures depicting the contrast to the focal meaning component. Similarly to Mulders and Szendrői (2016), we expected them to do so to verify the focal meaning component (that the dinosaur is doing something other to the ball than kicking it). To determine whether music training has a positive effect on the prosody-semantics interface, we then analysed if more music training correlated with more fixation on the picture depicting the contrast to the focal meaning component. Based on previous literature, we hypothesised that participants who had received more music training or participants who actively listened to music more often would be more aware of the prosodic information (pitch accents) in our experimental utterances. As

a result, we predicted that they would be better or faster at processing this information and integrating it into the semantics of the utterances (more details on our predictions in section 3).

2. Method

This study was conducted as part of a larger study on the processing of prosodic information in sentences in English with 'only' (Ge et al., in progress). An eye-tracking experiment was conducted on Dutch learners of English. The visual world paradigm was used to test our hypothesis. This means that the participants listened to auditory stimuli and their eye-movements were tracked to determine how the auditory stimuli influenced their interaction with our visual stimuli. For the purpose of the current study, we obtained detailed information on the music training our participants had received and on the amount of music they actively listen to per week.

2.1 Participants

In total, 34 participants participated in the experiment. Of the 34 participants, 31 successfully completed the eye-tracking experiment. One participant was also excluded because they did not include enough information on their musical history. The participants were *English language and culture* students from Utrecht University and Leiden University with Dutch as their native language. The students had to sign up to participate and were paid 10 euros for participating. The ages of the participants ranged from 17 to 29 years old with a mean age of 21.33 years old. In total, 19 participants reported having had some form of music lessons.

Based on previous research (cf. Schellenberg & Moreno, 2009; Wayland et al., 2010), we determined the total number of years that the participants had received music training. Years of practice without music lessons were not counted towards this number. The number of lessons for the 19 participants who received music lessons ranged from 1 to 21 years with a

mean of 8 years. Nearly all the participants reported having had training on piano (N=15). For more detailed information on the music training the participants received, see Appendix A. To discover if listening to music often has a similar effect as formal music training, the participants were asked how many hours per week they actively listened to music. On the questionnaire, actively listening to music was described as 'consciously listening to music and its different components (lyrics, composition), singing/humming along, etc.'. The participants were also asked how many hours per week they listened to music passively (which was described as 'music in the background while focusing on other things') in an attempt to retrieve information on actively listening to music more accurately. The number of hours the participants reported actively listening to music ranged from 0 to 50 hours per week with a mean of 8.3 hours per week.

2.2 Materials

The participants listened to 48 short stories about four animals interacting with different objects. Four of these items were used as practice items to make the participants familiar with the procedure and materials. Twenty-four of the items were fillers and the 20 remaining items were the experimental items. An example of the experimental items can be seen in (6). All the items in our experiment followed a similar pattern. Animals were introduced along with some objects and what they potentially were going to do to them. After the participants were told that 'he/she changed his/her mind', the resolution followed. In the experimental items, the resolution to the story contained the focus particle 'only' and either the first verb or the first object ('verb1' and 'object1', respectively; see Figure 2) emphatically marked by a pitch accent to mark focus, as can be seen in (6b,c). Both the verb-condition and the object-condition contained 10 items. A male native speaker of British English recorded the audio for these items in a sound-proof room. The stimuli had a sampling frequency of 44.1kHz and a 16bit resolution. All items were Latin-squared and pseudo randomised.

6. a) The dinosaur had a bucket and a suitcase. He was going to carry them and throw them. Then he changed his mind.

b) The dinosaur is only CARRYING the bucket, not throwing the bucket.

c) The dinosaur is only carrying the BUCKET, not carrying the suitcase.

The auditory stimuli were accompanied by images depicting the possible resolutions of the different items. Each image had a resolution of 1024 by 768 pixels to match the dimensions of the screen in the UiL OTS eye-tracking labs. The images contained four pictures, depicting four possible resolutions to the short introductory story (Figure 1). Based on the example in (6), one of the pictures depicted the non-focal meaning component (Figure 1a) that the dinosaur is carrying the bucket. This picture is the target. One of the pictures depicted a possible contrast to the focal meaning component in the case that the resolution had focus marked with a pitch accent on the first object (= the dinosaur carrying the suitcase). This picture was the object-alternative picture (Figure 1b). A third picture depicted a possible contrast to the focal meaning component in the case the resolution had focus marked with a pitch accent on the first verb (= the dinosaur throwing the bucket). This picture was the verb-alternative picture (Figure 1c). Lastly, the remaining picture depicted the dinosaur doing the action that was not described in the first part of the resolution (= throwing) to the object that was not described in the first part of the resolution (= the bucket). This picture was the distractor picture. The spatial positions of the pictures varied in the experimental items but were balanced across conditions.

The participants were tested using the EyeLink 1000 in a sound treated room in the UiL OTS labs in Utrecht. The EyeLink 1000 system uses infrared based tracking technology to track one eye at a sampling rate of 500 Hz. Freedom of movement is 32 x 21 cm at a 60 cm

distance with a gaze position accuracy of 0.4 degrees. Through the use of a questionnaire, the participants were asked to provide several details about their musical background, among other things (see Appendix B for the full questionnaire).

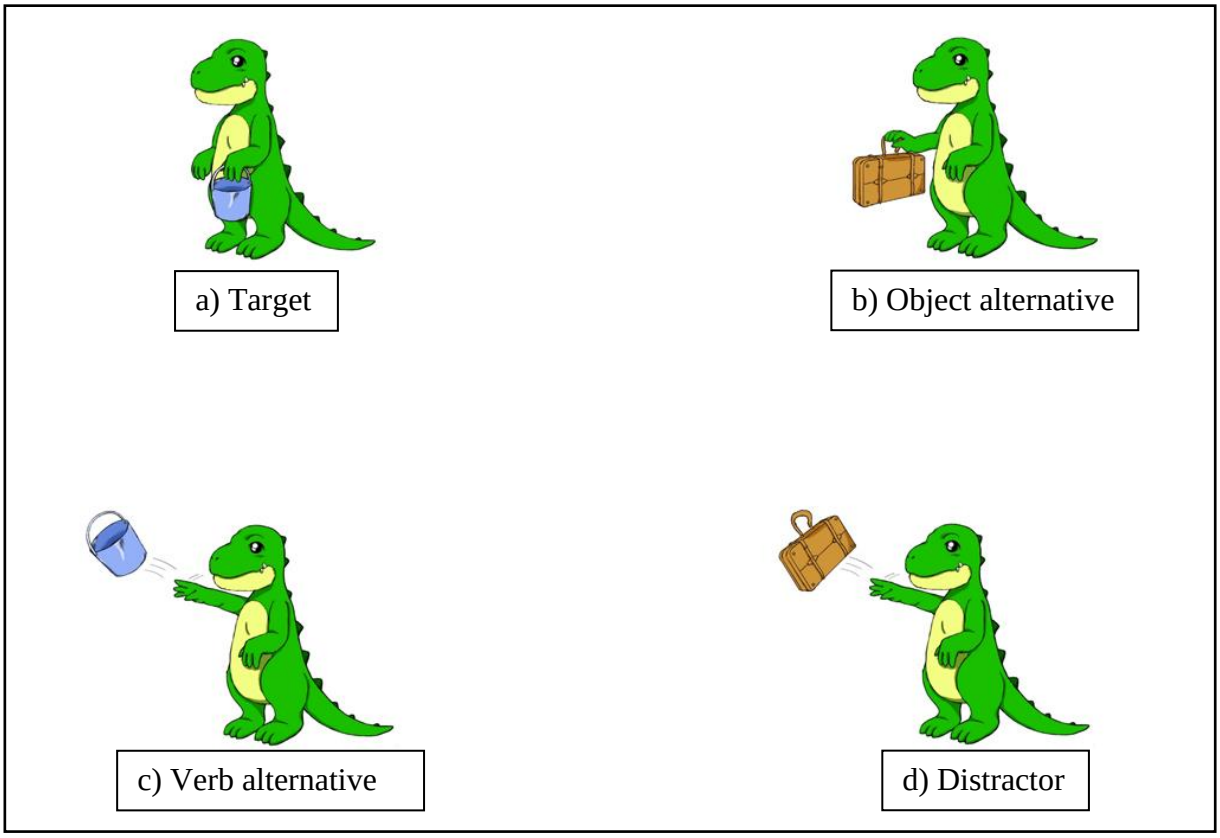

Figure 1. An example of the visual stimuli that were presented to the participants to accompany the example in (6).

2.3 Procedure

The participants were not informed of the design or purpose of the experiment beforehand. Instead, they were instructed to simply listen to the short stories they were about to hear and look at the images on the screen. The participants were asked to sit down in a chair in front of a monitor and the EyeLink 1000 camera. After permission, a target sticker was placed on the forehead of the participants to help the camera track the right eye movement. The participants were positioned between 550mm and 650mm from the eye-tracking camera. The height of the chair was adjusted to make sure the participant's eyes were

roughly in the centre of the camera's view. A 13-point calibration and validation procedure was used to ensure that the EyeLink 1000 was tracking eye movement correctly. The 4 trial items were then used to make the participants aware of the procedure and to ensure everything was set up properly. If needed, adjustments to volume and the eye-tracking set up were made before starting the experiment. The experiment lasted for roughly 20 minutes.

3. Predictions

We hypothesised that participants who received more music training or who actively listened to music more often than other participants would be better at integrating prosodic information (pitch accents) into the semantics of the utterances with the focus particle 'only'. More specifically we expected the participants to integrate prosodic information as soon as they heard it based on the findings by Mulders and Szendrői (2016). Based on the literature suggesting that music training can give musically trained participants advantages when it comes to perceiving and processing prosodic information, we also expected musically trained participants to perform better than participants who have received less music training. As explained, we expected participants to look more at the picture depicting a contrast to the focal meaning component when the focal meaning component became apparent through the use of a pitch accent on the focused constituent (either 'verb1' or 'object1', see Figure 2). Given the expectations that musically trained participants perceive and process prosodic information better, we predicted that participants who received more musical training would have higher proportions of fixations to the verb alternative picture when the 'verb1' was marked as the focused constituent by a pitch accent. We expected the same for the proportions of fixation on the object alternative picture when 'object1' was marked as the focused constituent by a pitch accent.

Based on evidence from the literature, there may be an effect from actively listening to music on musical abilities. To explore how actively listening to music can affect linguistic

and specifically prosodic abilities, we asked the participants how much they actively listened to music per week. Assuming that actively listening to music can help perceiving and processing prosodic information, we predicted that participants who actively listened to music more often than others would have higher proportions of fixations on the verb alternative picture when the 'verb1' was marked as the focused constituent by a pitch accent. We also expected this for the proportions of fixation on the object alternative picture when 'object1' was marked as the focused constituent by a pitch accent.

4. Results

4.1 Preliminary remarks

To analyse our results, we used multiple regression analyses to explore whether more musical training would correlate with a higher proportion of looks to the alternative picture. In most of the literature on the effects of music training, participants are usually divided to be or selected as part of musician or non-musician groups (cf. Schön, et al., 2004; Schellenberg & Moreno, 2009; Wayland et al., 2010). However, in Schellenberg and Moreno (2009), for instance, participants who had received a limited amount of music training were not counted towards the musically trained participants, since the training was not considered extensive enough. It seems that drawing the line between musicians and non-musicians can be somewhat arbitrary. Moreover, in the current study, we were not in a position to exclude participants who had not received enough music training due to limited sample size. Considering this, we decided it was best to not reduce years of music training to a categorical independent variable and instead keep it a gradual one. Therefore, we decided to use multiple regression analysis. Regarding listening to music and its effects, there is far less literature on this phenomena. Turning listening to music into a categorical independent variable was an equally difficult task. We therefore also treated it as a gradual variable in the multiple regression analysis.

There are three points at which we could compare proportions of fixations (Figure 2). Upon hearing ‘only’, the Dutch learners of English might have expected the focused component to follow directly since this is most common in Dutch (Dimitrova, 2012; Foolen et al., 2009). This means that the participants could be biased to expect a contrast in ‘verb1’. This would not be likely to lead to an increase in fixation on any of the pictures, since the participants could not yet have known which verb would be accented. The focus particle ‘only’ is the reason that a pitch accent is needed to understand what the focal meaning component is and to resolve ambiguity about which of the two possible focal meaning components could be part of the sentence. Upon hearing ‘only’, the participants had likely not yet received enough information to know what the non-focal or focal meaning component could have been. Therefore, each of the four pictures was a possible contrast to the possible focal meaning component. Because of this, the first window of interest started at the onset of ‘only’ and lasted until the onset of ‘verb1’ (Figure 2). This window could be analysed to make sure there is no effect of music training or listening to music before any of the benefits could play a role. The proportion of fixation for the target picture should be left out of the analysis since this picture played no further role in the analysis of the current study. Furthermore, the placement of pitch accent is not relevant as an independent variable, since there is no difference between the verb accented or the object accented utterances, until after ‘verb1’.

From the onset of ‘verb1’ until the onset of ‘object1’ was the first moment where, with a pitch accent on ‘verb1’, we could have found an increase in the proportion of fixation on the verb alternative picture. It should be noted at this stage that before the participants hear ‘object1’ in the utterance, there are two possible verb-alternative pictures in the image. Before hearing that ‘object1’ is the ‘bucket’, the image depicting the dinosaur throwing the suitcase is also a verb alternative picture, because both the distractor and the verb-alternative picture depict the dinosaur throwing an object instead of carrying it. Therefore, the proportions for

both the distractor and verb alternative image are taken together until ‘object1’ has been introduced. This was window 2. At this point, music training and listening to music could have started to play a role, as well, since this was the first moment that prosodic information needed to be processed to understand the semantics of the utterance. The participants were likely to look at the alternative picture once they were aware of the focal meaning component. At the onset of ‘object1’, there was another window where we expected to find an increase in the proportion of fixation, in this case for the object alternative picture in combination with a pitch accent on ‘object1’. Window 3 lasted until the onset of ‘verb2’.

Of these three windows, we only statistically analysed window 3 for the items with a pitch accent on ‘verb1’. Multiple regression analysis was the best option because of our gradual music variables, but it should not be used with repeated measures. Therefore, we decided to use the multiple regression analysis on only one window of interest. The difference in proportion of fixation based on the placement of the pitch accent was the biggest for the object-alternative picture during window 3 (as can be seen in Figure 3). Moreover, the proportion of fixation on the object-alternative picture during window 3 was largest for the verb-accented condition. Because of these considerations, window 3 in the verb-accented condition seemed like the best opportunity to observe potential effects of music training and listening to music on the processing of the prosody-semantics interface.

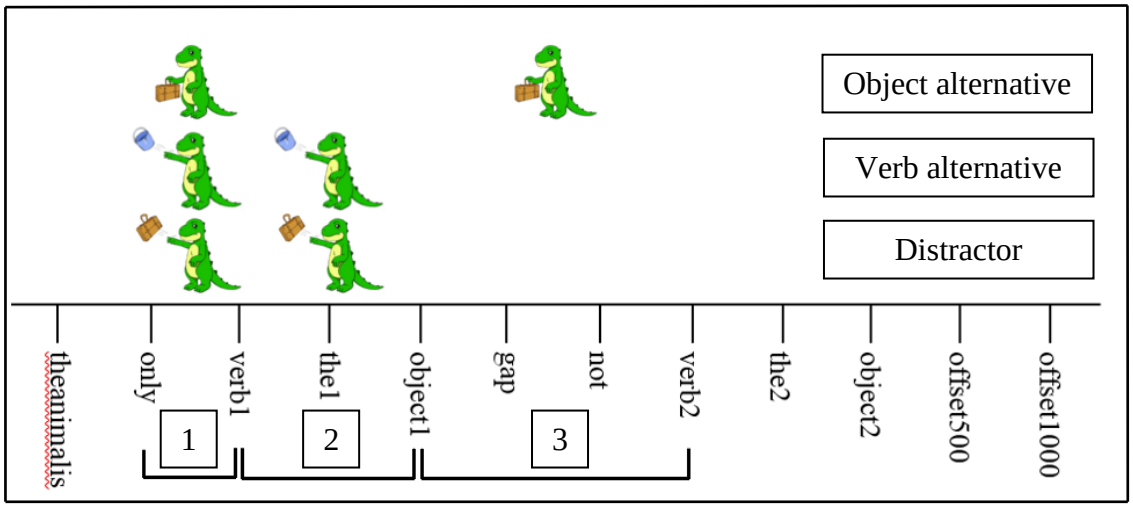

Figure 2. Diagram showing which pictures were part of the windows 1, 2 and 3.

4.2 Multiple linear regression

Mean proportions of fixation were calculated for the four pictures that the participants were presented with. The proportions were measured at the different parts of the sentences and for the verb-accented condition and object-accented condition separately. These mean proportions of fixation can be found in Figure 3. Multiple linear regression was calculated for window 3 in the verb-accented condition to predict the proportion of fixations on the selected images based on the amount of music training and active listening. In Table 1, the summary of the multiple linear regression analysis for window 3 can be found. No significant regression equation was found (F(2, 27) = 0.021 , p = .979) with an R^2 of .002. This indicates that 0.2 percent of the variance in the percentage of fixations can be accounted for by the current model. The predicted percentage of fixation for window 3 is equal to 34.068 - 0.028 + 0.016, where music training was measured in years and music listening in hours per week. However, neither years of music training, nor listening to music were significant predictors of the percentage of fixations on the object alternative image.

Table 1.

Summary of Multiple Linear Regression Analyses for Variables Predicting Proportion of Fixation for Window 3 (N = 30)

	B	SE	β
Constant	34.068	3.124	
Years of Lessons	- 0.028	0.319	- .017
Music per Week	0.016	0.090	.034

R^2 = .002 for Window 3 (p = .979)

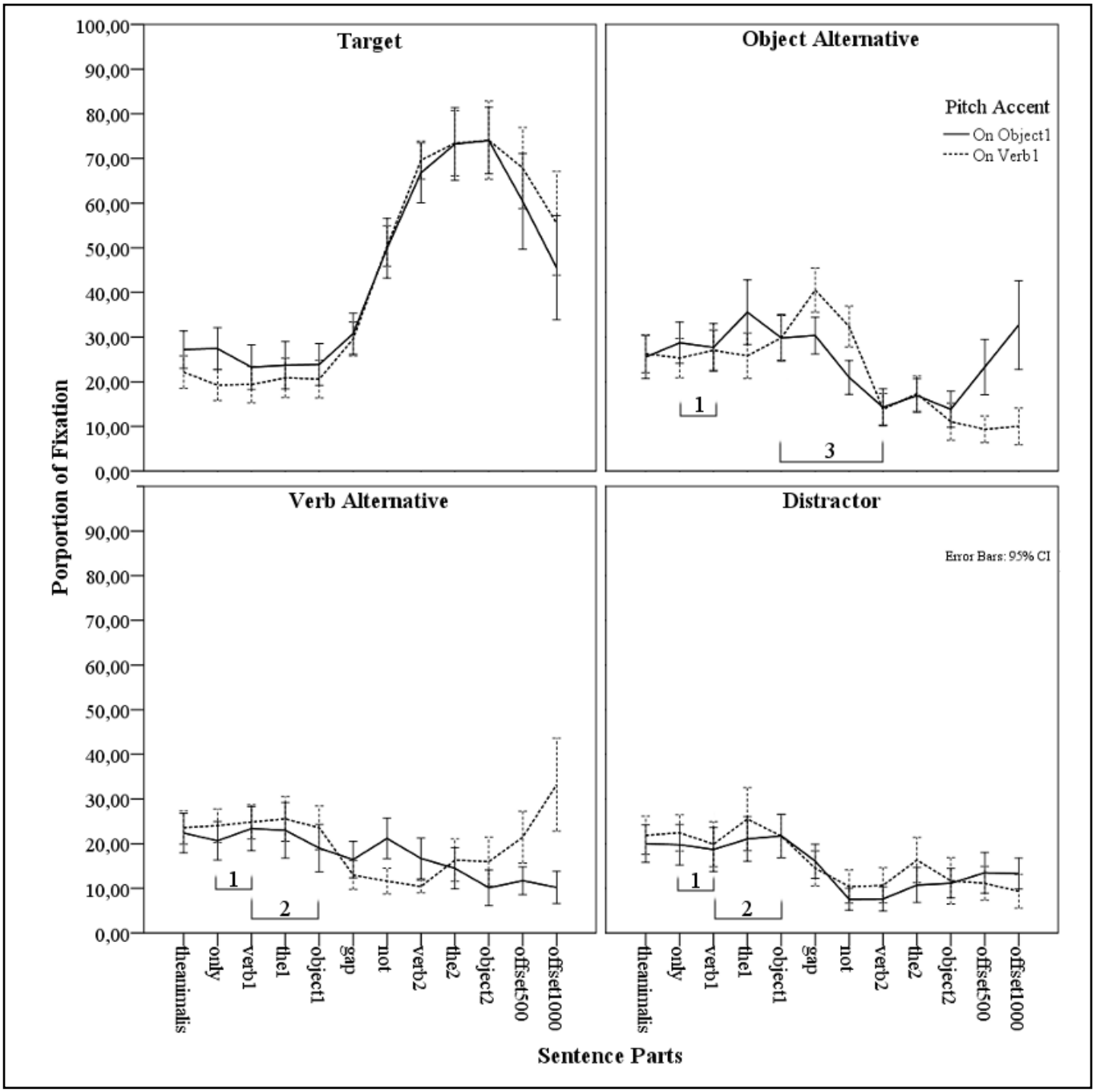

Figure 3. Mean proportion of fixation of all participants on the four different pictures for the experimental items with pitch accent on ‘verb1’ or on ‘object1’.

5. Discussion

The results of the regression analysis on window 3 were not as expected. No significant results were found for any of the entered variables into our model. This means that neither the amount of music training, nor the amount of music listening were good predictors

for the proportion of fixation on the object-alternative picture during window 3 when the participants heard resolution sentences with pitch accents on ‘verb1’. As can be seen in the proportion of fixation on the object-alternative picture in Figure 3, the increase in the proportion of fixation occurred when the utterances had pitch accents on ‘verb1’, and not on ‘object1’. Based on the literature, we expected the participants to increasingly look at the object alternative picture after hearing a pitch accent on ‘object1’ to confirm that the dinosaur was not carrying another object (e.g. the ‘suitcase’), since that would contrast with the focal meaning component of the resolution that the participants heard in the sentence with a pitch accent on ‘object1’. However interesting and somewhat surprising this result is, the main focus of this paper was on the effect of music training on the processing of prosodic information[1]. Ultimately, the results of the regression analysis for window 3 suggest that neither music training nor actively listening to music gave participants an advantage in perceiving and processing the prosody-semantics interface.

These results, then, contradict our predictions as well as a large portion of the literature discussed in the theoretical framework. The question is why this happened. There is a reasonable chance that the results we have found contradict predictions and literature because of methodological issues. While 30 participants in some cases can be a sufficient sample size, when trying to explore a relatively small effect of music training, there is a chance that this sample was simply not large enough to detect an effect of music training or listening to music. Especially with regards to the multiple regression analysis, Fields (2009) suggested that with three predictors in a model where a medium or small effect is expected, a sample size of between 80 to over 500 participants might be needed for the analysis to have enough power. This shows that especially for the multiple regression analysis, more participants were needed to find meaningful results.

[1] Ge et al. (in progress) will look more at the findings regarding the pitch accents

Furthermore, there are some methodological issues regarding music training to be considered. Patel (2014) proposed in his OPERA hypothesis that the effects of music training on speech processing are partially due to the frequent repetition that comes with music training. While we did have a pretty good understanding of the formal training our participants received, we may not have had enough insight into their personal music habits with regards to self-practice. If repetition is an important aspect of why and how music training affects speech perception, it might be crucial to have a better understanding of how intensively music has been practiced by the participants. Moreover, in the current study, we made no distinction between music training for different disciplines. In exploratory fashion, Lima and Castro (2011) found that there was no significant difference between different instruments in their experiment on the effect of music training on the recognition of emotions in speech prosody. However, Thompson et al. (2004) suggested that there is a difference between instrumental and vocal music training, as some aspects of vocal training may interfere negatively with the understanding of prosodic uses of the voice in speech. Since we made no distinction between instrumentalists and vocalists, this could have also affected our findings.

Besides methodological considerations, there are also other potential reasons that could explain the results we found. Current literature (Gennari, 2005; Mulders & Szendrői, 2016) has shown that Dutch and English speakers are sensitive to focus cues in their respective languages and more specifically, Mulders and Szendrői (2016) showed that Dutch speakers use prosodic information to interpret the semantics of utterances. However, there appear to be some structural differences between Dutch and English in relation to ‘only’ and its focused element (Bouma et al., 2007; Foolen et al., 2009). Foolen et al. (2009) showed that in Dutch, the focused element often follows ‘only’ immediately. In English, however, ‘only’ is often still placed before the verb even when the object receives focus. Perhaps the Dutch

learners of English were aware of this difference and were, therefore, slower with integrating the prosodic information to wait until they heard the other potentially focused element. Even so, this would not explain why we found the results for window 3. It also does not explain why music training did not help the musically trained participants to perceive and process the prosodic information better than the less musically trained participants. More research is necessary to explain the current findings.

In an attempt to explain why the amount of music training participants had received was not a good predictor of the proportion of fixation, we should take another look at the literature. For instance, it has been suggested that there are factors other than music training, such as emotional intelligence, that are better at predicting how perceptive participants are to emotions based on speech prosody (Trimmer & Cuddy, 2008). This suggests that other factors than music training could explain why certain participants outperform others. Similarly, the link between music training and intelligence has recently been revisited. Swaminathan, Schellenberg, & Khalil (2017) found that testing participants on music aptitude was a better indicator than music training. Perhaps the same is true for the link between music training and linguistic ability, and specifically prosodic ability. It could be that we should have looked at the music aptitude of the participants and that this is partially why we did not find the results we expected based on the literature.

Finally, it seems that the results also contradicted our prediction in regards to actively listening to music. Since this variable was considered explorative, there is little literature to explain why did not find the results we expected. Perhaps, we found no effect from listening to music because of similar methodological issues. It could also be that we found no effect from listening to music because listening to music often does not help participants perceive and process prosodic information, but more research is necessary to confidently determine these suggestions.

6. Conclusions and Future Research

In conclusion, the current study was unable to find a link between music training and the workings of the prosody-semantics interface in Dutch speakers of L2 English. No correlation was found between the amount of music training nor the amount of listening to music and how the participants performed in the eye tracking experiment. Some of these findings may be a result of methodological issues which should be considered with future research. However, it could be that the link between prosodic processing and perception and music needs to be explored in more detail since our results suggest that this connection may not be as strong or well defined as previous research suggests.

Based on our findings and considerations, there are many possible directions for future research. Firstly, a similar experiment with more participants could be more efficient at exposing the possible effect of music training on the processing of prosodic cues. If a similar experiment were to be conducted, a better and more in depth understanding of the participants' habits in regards to music training would be beneficial, as well. Besides attempting to improve on the current study, there is also an opportunity to explore the benefits of music training some more. Perhaps exploring the link between music aptitude and prosodic perception and processing turns out to be more robust than the link between music training and prosodic benefits.

References

Bigand, E., & Poulin-Charronnat, B. (2006). Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. *Cognition,100*(1), 100-130.

Bouma, G., Hendriks, P., & Hoeksema, J. (2007). Focus Particles Inside Prepositional Phrases: A Comparison of Dutch, English, and German. *The Journal of Comparative Germanic Linguistics,10*(1), 1-24.

Dimitrova, D. V. (2012). *Neural Correlates of Prosody and Information Structure.* (Ph.D. Thesis). Retrieved from Rijksuniversiteit Groningen.

Dimitrova, D. V., Stowe, L. A., Redeker, G., & Hoeks, J. C. J. (2010). Focus particles and prosody processing in Dutch: Evidence from ERPs. *Speech Prosody 2010*, 1-4.

Foolen, A., Van Gerrevink, R., Hogeweg, L., & Prawiro-Atmodjo, P. (2009). The placement of focus particles in Dutch. *Linguistics in the Netherlands,26*(1), 51-63.

Gennari, S. P. (2005). Rapid relief of stress in dealing with ambiguity. In J. Trueswell & M. Tanenhaus (Eds.), *Approaches to studying world-situated language use* (pp. 245-259). Cambridge: MIT Press.

Gualmini, A., Maciukaite, S., and Crain, S. (2003) Children's insensitivity to contrastive stress in sentences with only. *University of Pennsylvania Working Papers in Linguistics,8*(1), 87-100.

Honing, H., & Ladinig, O. (2009). Exposure influences expressive timing judgments in music. *Journal of Experimental Psychology: Human Perception and Performance,35*(1), 281-288.

Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion,11*(5), 1021-1031.

Marques, C., Moreno, S., Castro, S. L., & Besson, M. (2007). Musicians Detect Pitch

Violation in a Foreign Language Better Than Nonmusicians: Behavioral and Electrophysiological Evidence. *Journal of Cognitive Neuroscience,19*(9), 1453-1463.

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009) Musical Training Influences Linguistic Abilities in 8-Year-Old Children: More Evidence for Brain Plasticity. *Cerebral Cortex,19*, 712-723.

Mulders, I., & Szendrői, K. (2016). Early Association of Prosodic Focus with alleen 'only': Evidence from Eye Movements in the Visual-World Paradigm. *Frontiers in Psychology,7*, 1-19.

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The Musicality of Non Musicians: An Index for Assessing Musical Sophistication in the General Population. *PLoS ONE,9*(2), 1-23.

Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research,308*, 98-108.

Rauscher, F. H., & Hinton, S. C. (2006). The Mozart Effect: Music Listening is Not Music Instruction. *Educational Psychologist,41*(4), 233-238.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics,1*(1), 75-116.

Schellenberg, E. G., & Moreno, S. (2009). Music lessons, pitch processing, and g. *Psychology of Music,38*(2), 209-221. doi:10.1177/0305735609339473

Schellenberg, G., & Weiss, M. W. (2013). Music and Cognitive Abilities. In D. Deutsch, *The psychology of music* (pp. 499-550). Amsterdam: Elsevier, Acad. Press.

Schirmer, A., Kotz, S. A., & Friederici, A. D. (2002). Sex differentiates the role of emotional prosody during word processing. *Cognitive Brain Research,14*(2), 228-233.

Schirmer, A., Zysset, S., Kotz, S. A., & Cramon, D. Y. (2004). Gender differences in the activation of inferior frontal cortex during emotional speech perception.

NeuroImage,21(3), 1114-1123.

Schön, D., Magne, C., & Besson, M. (2004) The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, 41, 341-349.

Swaminathan, S., & Gopinath, J. K. (2013). Music training and second language English comprehension and vocabulary skills in Indian children. Psychological Studies, 58, 164–170.

Swaminathan, S., Schellenberg, E. G., & Khalil, S. (2017). Revisiting the association between music lessons and intelligence: Training effects or music aptitude? *Intelligence,62*, 119-124.

Thompson, W. F., Schellenberg, G., & Ilie, G. (2004). Perceiving prosody in speech. Effects of music lessons. *Annals of the New York Academy of Sciences,4*(1), 46-64.

Trimmer, C. G., & Cuddy, L. L. (2008). Emotional intelligence, not music training, predicts recognition of emotional speech prosody. *Emotion,8*(6), 838-849.

Wayland, R., Herrera, E., & Kaan, E. (2010). Effects of musical experience and training on pitch contour perception. *Journal of Phonetics,38*(4), 654-662.

Appendix A. *Information on the music training the participants have received and how many hours they actively listen to music per week*

Participant number	Age	Sex	Age range of music lessons	Total years of music lessons	Currently practicing	Music Discipline 1	Music Discipline 2	Music Discipline 3	Music Discipline 4	Active music listening p/w
pp101	21	F	6-21	21	No	Vocals	Piano	Guitar		6
pp102	24	M	9-22	12	Yes	Flute	Piano	Drums	Saxophone	10
pp103	21	F		0	No					1
pp104	27	M	8-17	9	No	Piano				2
pp106	29	M		0	No					1
pp107	27	M		0	No					10
pp109	20	M	13-17	4	No	Piano				2
pp110	22	F	11-14	3	No	Piano				20
pp111	21	M		0	No					4
pp112	22	F	12-17	5	No	Guitar				10
pp113	21	M	10-12	2	No	Keyboard				20
pp114	20	M		0	No					4
pp115	20	M	8-18	15	Yes	Piano	Singing			2
pp117	20	F	10-15	5	Yes	Piano				5
pp118	22	F	9-18	8	Yes	Singing	Guitar	Piano		2
pp120	19	F	8-18	16	No	Cello	singing			5
pp121	17	F	7-11	3	No	Blokfluit	singing	violin		3
pp122	19	F		0	No					0
pp123	20	F	8-17	8	No	flute	guitar	piano		2
pp125	19	F	6-17	14	Yes	Piano	Guitar			2
pp126	21	F	7-8	1	No	Flute				12
pp127	20	F		0	No					30
pp128	21	F		0	No					2
pp129	19	F		0	No					7
pp130	20	M	6-17	8	Yes	Piano	Keyboard	Guitar	Vocals	50

pp131	21	M		0	No		2
pp132	22	F	12-16	4	No	Singing	10
pp133	21	F		0	No		4
pp134	22	M	8-18	10	Yes	Piano	7
pp135	22	F	8-12	4	No	Piano	14

Appendix B. *Questionnaire that participants were asked to fill out at the end of the experiment*

Language Background Questionnaire

Please answer all the following questions to the best of your ability. If a particular question does not apply to you, please fill in the appropriate space with an N/A.

1. Name: _____ 2. Age: _____ 3. Gender:_____
4. Country of birth: _____ 5. Native Language:

6. Rate your current **overall** language ability in **English**.
a. Native b. Excellent c. Good

7. Your current **English speaking** ability:
a. Native
b. Very fluent, i.e., can carry out virtually any kind of conversation in almost any situation
c. Quite fluent, i.e., can carry out some extended conversations and communicate effectively

8. Your current **English comprehension** ability:
a. Native
b. Excellent understanding, i.e., can understand almost everything in almost every situation
c. Good understanding, i.e., can follow extended conversations e.g., can follow films/TV shows

9. VWO English grade (and IELTS score, if applicable):_____

10. Please list **all** of the languages other than your native language you know, **in order in which you acquired/learned them from earliest to most recent**.

Non-native Language	Age of acquisition	Years of learning/using the language	Source of exposure (xxx, xxx, xxx)
1.			
2.			
3.			
4.			

11. Are you left or right handed?
a. Left b. Right c. Both

12. Do you currently suffer any neurological or psychological disorders?
a. Yes b. No
13. Were you diagnosed for any neurological or psychological disorders in the past?
a. Yes b. No
14. Do you experience trouble with your vision?
a. Yes b. Yes, but it's corrected (with glasses or lenses) c. No
15. Have you ever practiced any music instruments/vocals?
a. Yes b. No (skip question 16)

16. Please specify your musical training using the follow table:

Discipline (e.g. Guitar, opara)	Age range in which you received different types of training	Age range of practice if you no longer practice it currently	Hours of practice per week if you still practice it
	Primary school lessons:		
	High school lessons:		
	Private lessons:		
	Primary school lessons:		
	High school lessons:		
	Private lessons:		
	Primary school lessons:		
	High school lessons:		
	Private lessons:		
	Primary school lessons:		
	High school lessons:		
	Private lessons:		

17. How many hours a week do you listen to music passively (i.e. Music in the background while focusing on other things)? _____

18. How many hours a week do you listen to music actively? (i.e. consciously listening to music and its different components (lyrics, composition), singing/humming along, etc.) _____

19. What type of music do you mostly listen to?

a. Instrumental b. Classical c. Pop d. Rock/metal e. Hip hop/rap

Thank you very much filling in the questionnaire!