NON-NATIVE MEMORY TRACES CAN BE FURTHER STRENGTHENED BY SHORT TERM PHONETIC TRAINING

Henna Tamminen and Maija S. Peltola

Department of Phonetics, University of Turku, Finland
Learning, Age and Bilingualism laboratory (LAB-lab), University of Turku, Finland
Centre for Cognitive Neuroscience, University of Turku, Finland
henna.tamminen@utu.fi

ABSTRACT

Learners of a second language go through many proficiency levels during language learning. University language students can be considered advanced learners, who behaviourally perceive the language at least nearly similarly as native speakers. It can also be assumed that memory traces for the second language phonemes have developed for these learners and their MMN responses for foreign contrasts may be even native-like. Here, we wanted to find out, whether these students have memory traces for second language speech categories. Further, we investigated whether their behavioural perception can be improved and can the memory traces be strengthened by training. For this purpose we tested Finnish university students of English on a three-day listen-and-repeat training procedure. The baseline MMN response was further strengthened and category boundary became more consistent, reaction times decreased and discrimination sensitivity increased. The perception – behavioural and psychophysical – of the learners of English was further strengthened by listen-and-repeat training.

Keywords: speech perception, phonological processing, training, mismatch negativity (MMN)

1. INTRODUCTION

By the age of six months speech sounds are perceived according to the mother tongue system [4] since the native language memory traces develop in early childhood [1]. At least at the initial stages of foreign language learning, the non-native language is perceived through language specific native language memory traces [6].

The learning of foreign language perceptual patterns and the development of memory traces for the foreign phonemes has been studied, for example, by Peltola et al. [8] where it was shown that Finnish students of English do not have native-like mismatch negativity (MMN) responses for target language categories and their response to the mother tongue is smaller than Finnish monolinguals' response to the same contrast. They suggested that this could implicate incompleteness of learning and that the two language systems may be intertwined. From another point of view, the linguistic context has an effect on learners' perception and it seems that there are two independent phonological systems in language learners so that the contextually redundant system can be switched off [9, 10]. It is clear that both the different stages of learning and the different linguistic contexts have an impact on second language perception. And surely, the role of the mother tongue is different in different phases of learning.

Learning effects and neural plastic changes have also been shown in training studies. For example, Menning et al., [5] showed, using both behavioural and MEG testing, that after training German learners were able to discriminate Japanese words. Also, Tamminen et al., [11] showed that native Finns learn to perceive a phonemically irrelevant voicing contrast and that through training, native-like memory traces evolve in three days for this difficult acoustic feature.

In this study, we wanted to see whether a three-day listen-and-repeat training affects language learners' perception at the behavioural and psychophysiological levels. We studied this with behavioural methods using identification (ID), goodness rating (GR), reaction time (RT), and discrimination tests together with psychophysiological MMN recordings. Since the language learners master the target language well, the hypothesis was that the baseline results would be different from the native Finns, who do not master the trained contrast, but who showed perceptual changes during the whole three-day period at both attentive and pre-attentive levels [11]. Due to high non-native proficiency, the learners should perceive the contrast and be able to identify and rate the goodness of the category members presented already prior to any training. In addition, they were expected to both behaviourally and neurally discriminate the contrast and to perhaps already have a memory trace for the target contrast before training. Both perceptual changes and strengthening of memory traces were expected in the advanced learners due to the listen-and-repeat training. Since the students were different from the native Finns to start with, the progress in training was considered to be quite different. For that reason, the most interesting comparison was that between the baseline measurement and the final registration after the training protocol. It may also be that the effects are seen only in some levels of perception or the effects may be shown simultaneously at the behavioural and the pre-attentive levels as was reported by Tremblay et al. [13].

2. METHODS

2.1. Subjects and stimuli

The subjects were 11 native Finnish students of English (age range 20–28, mean age 23.6, 6 females). They were English majors who had studied English for 3.5 years on average at university level (range 1–7 years) and 10 years before that in primary and secondary schools. Subjects were right-handed (tested with Edinburgh Handedness Inventory [7]) and they had normal hearing (tested with an audiometer). Prior to participating in the experiments, an informed consent was obtained from each subject.

The stimuli consisted of 15 synthesised (HLSyn software 1.0 Sensimetrics Inc.) variants of the English words /fi:l/ *'feel'* and /vi:l/ *'veal'*. The voice onset time (VOT) of the first sound was varied in 14 ms steps; the other end of the continuum was totally voiceless and the other completely voiced. The 499 ms long stimuli were identical from 197 ms onward, i.e., during the vowel and lateral part of the stimuli. The whole continuum was used in the ID and GR tests. Two words from the continuum were selected for the training, discrimination and RT experiment and MMN registrations. These words had been classified as belonging to different categories by native English speakers in an earlier ID test. The VOTs of these stimuli were 113 ms and 71 ms for /fi:l/ and /vi:l/, respectively. Therefore, the difference between the stimuli started at 71 ms.

2.2. Procedure and analysis

In the forced choice ID and GR experiment stimuli were presented 8 times in random order. The subjects were instructed to label the stimuli as 'feel' or 'veal' and rate the goodness of the word on a 1–7 scale in which 1 was poor and 7 excellent.

In the discrimination and RT experiment the stimuli were presented in an oddball paradigm where /fi:l/ was the standard and /vi:l/ the deviant. Deviant probability was 0.13 (130 standards and 20 deviants) and the inter stimulus interval (ISI) was 1000 ms. The subjects were instructed to press the response button as soon as they heard the deviant.

The listen-and-repeat training was self-paced and 60 stimuli were presented so that every other one was /fi:l/ and every other /vi:l/. The instructions were to listen to and then repeat each stimulus carefully.

The stimuli in the MMN registrations were also presented in the oddball paradigm while the subjects watched a silent non-subtitled movie of their choice. Deviant probability was again 0.13 (783 standards and 120 deviants) and ISI was 650 ms. The EEG was recorded with 21 Sn electrodes (Electro-Cap International, Inc.) and Synamps amplifier (sampling rate 250 Hz; bandwidth 0.5–70 Hz). Electrodes attached below and near the outer canthus of the right eye monitored eye movements. Impedance was kept under 5 kΩ.

The three day study protocol was as follows:

- ID, GR, discrimination, RT, MMN, and training
- training, ID, GR, discrimination, RT, MMN, and training
- training, ID, GR, discrimination, RT, and MMN

The order of the behavioural discrimination and RT experiment and the MMN registration was counterbalanced. The subjects received no feedback during any of the experiments. The study was conducted in accordance with the guidelines of the Ethics Committee of the University of Turku.

The category boundary location and steepness values were obtained from the logit transformation analysis (SPSS). Category boundary is the point where there is a 50% distribution of answers and the consistency of the subject's answers is indicated by the steepness value. The boundary location and steepness data were separately subjected to a Repeated Measures Analysis of Variance (ANOVA) between the two Days. The GR data was analysed for five stimuli with a Day (2) × Stimulus (5) ANOVA. The five stimuli in this analysis were the boundary, training stimuli, and one stimulus from both categories with the best rating in the first Day. The discrimination sensitivity (*d'*) was calculated according to the hits, misses, false alarms, and correct rejections. The RTs were measured from the onset of the deviant stimulus and the answers within ±3 standard deviation were included in the analysis. RT and *d'* values were separately subjected to a Repeated Measures ANOVA.

The EEG was off-line filtered with a 1–30 Hz bandpass filter and the artefact rejection criterion was set at ±100μV. The analysis epochs started 50 ms before and ended 500 ms after stimulus onset. Baseline was corrected during the 50 ms pre-stimulus period. Epochs were averaged for standards and deviants separately, and the response elicited by the standard stimulus was subtracted from the deviant response to create a difference waveform. For the mean amplitude analysis we selected two imbricated 40 ms time windows around the maximum amplitudes in the difference waveforms (Day 1: 300–340 ms, Day 3: 320–360 ms) and for the latency analysis a 160 ms time window (250–410 ms). Electrodes Fz, Cz, F3, F4, C3, C4 were selected for the analysis of the mean amplitude and Cz for the latency analysis. First, we tested whether the MMN response significantly differed from zero. Then, a Day (2) × Electrode (6) ANOVA was used in the mean amplitude analysis and the latency analysis was carried out between the two Days.

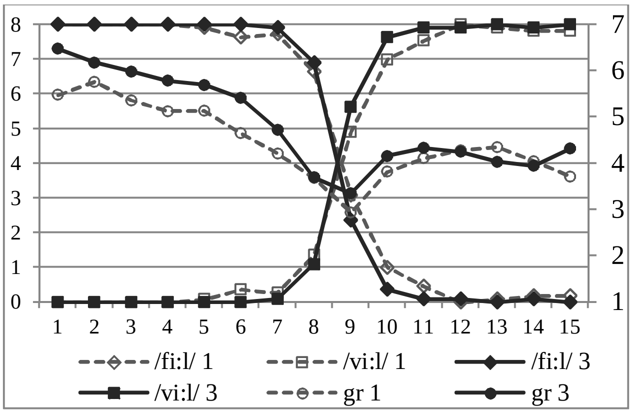
Post hoc tests were used where appropriate. Since the students master the target language so well it can be assumed that an MMN response elicits already during the baseline measurement. Therefore, the possible training induced enhancement of the response is expected to require the maximum amount of input. All analyses were thus performed within or between the baseline and the last day of training and testing. The whole training and testing protocol was, however, conducted in accordance with our earlier study in order them to be comparable and to have the same amount of training.

3. RESULTS

The ID results (Fig. 1) showed that category boundary steepness changed significantly ($F(1,10) = 6.652$, $p = 0.027$) but that there was no change in the boundary location as a result of the training. The steepness was 1.6 (std dev 0.5) in the baseline measurement and 2.0 (std dev 0.3) in the last day. The baseline category boundary was at 8.9 (std dev 0.6) and the final at 8.7 (std dev 0.4) on the third day.

Statistical analysis revealed no significant changes in the GR results between the baseline and the final testing since there was no main effect of Day. The interaction between the goodness rating and the Day did not reach significance either. However, the main effect of the goodness ratings of the five selected stimuli was found ($F(4,7) = 30.497$, $p < 0.001$) suggesting that different stimuli were rated differently. Further Paired Samples Test showed that all stimulus pairs, except for 7-10 and 7-13 on both Days and pair 10-13 on the last Day were significantly different. These results are shown in detail in Table 1. The GR results are also shown in Fig. 1.

Figure 1: ID and GR scores from Day 1 (dash line) and 3 (continuous line). X-axis shows the continuum (1=completely voiceless). Left y-axis shows the number of times (max 8) stimuli were identified as members of each category. The GR scale was 1–7 (1=poor, 7=excellent; right y-axis).



The discrimination results showed that the RTs decreased ($F(1,10) = 5.938$, $p = 0.035$) and the discrimination sensitivity improved ($F(1,10) = 10.627$, $p = 0.009$) during training. The RT was 674 ms (std dev 98) in the baseline measurement and 608 ms (std dev 130) in the final measure. The *d'* was 3.6 (0.8) and 4.4 (0.4) in the first and last measurements, respectively.

The MMN responses were existent at Fz, Cz, F4, and C4 electrodes in the baseline measurement and in all electrode positions in the final measurement since they significantly differed from zero ($p < 0.05$ in Day 1 and $p < 0.001$ in Day 3). The MMN mean amplitude analysis showed the main effect of Day ($F(1,10) = 12.186$, $p = 0.006$) suggesting that the MMN amplitude was larger on the third Day compared to baseline. The MMN results are shown in Fig. 2 and Table 2. The MMN peak latency was not significantly different on the two Days. The latency of the peak amplitude in the Cz electrode was 321 ms on the first day and 334 ms on the third Day.

Table 1: Paired samples t-test results for GR stimuli 2, 7, 9, 10, and 13 on Days 1 and 3. Stimuli 2 and 13 are the best rated category representatives, stimuli 7 and 10 are the training stimuli, and stimulus 9 is the boundary.

Stim pairs		2/7	2/9	2/10	2/13	7/9	7/10	7/13	9/10	9/13	10/13
Day 1	p	<.001	<.001	<.001	=.002	=.009	=.280	=.703	=.001	<.001	=.002
	t	5.705	8.441	6.116	4.154	3.235	1.142	-.392	-4.341	-7.416	-4.054
Day 3	p	<.001	<.001	=.001	=.001	=.005	=.253	=.173	=.003	=.047	=.644
	t	7.590	8.189	4.767	4.675	3.605	1.212	1.468	-3.957	-2.267	.477

Figure 2: MMN results for Days 1 and 3. The difference between the stimuli started at 71 ms.

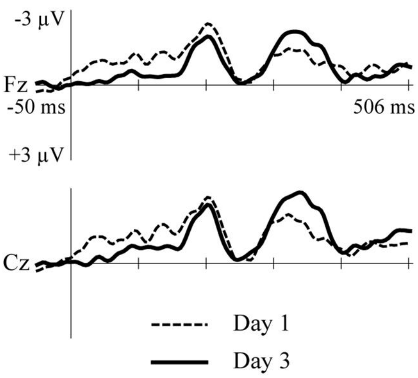


Table 2: Mean amplitudes (standard deviations in brackets) for six electrodes in Days 1 and 3.

	Fz	Cz	F3	F4	C3	C4
D1	-1.34	-1.78	-1.09	-1.27	-1.30	-1.53
	(1.96)	(1.93)	(2.03)	(1.79)	(1.97)	(1.48)
D3	-1.97	-2.62	-1.66	-2.28	-1.94	-2.44
	(0.86)	(1.50)	(0.99)	(1.20)	(1.17)	(1.26)

4. DISCUSSION AND CONCLUSIONS

Highly advanced university students have already high command of the target language. Here, we tested whether these students' perception of the target language can be changed by training. Similarly, as Iverson et al. [3] showed training effects on experienced language learners, our results showed that listen-and-repeat training has effects on the students' perception. Continuing mother tongue use has been shown to hinder the learning of second language perception [2] which may very well be a factor here as well. The results can be explained from three points of view, namely, mother tongue, prior target language skills and training.

Both, prior target language skills and training can be seen in the ID results. The subjects placed the category boundary between the training stimuli already at the beginning and so the training did not have an effect on the prior target language skills. However, the steepness of the boundary did change since it became more systematic. In other words, the advanced level of English was shown in the boundary placement, but training strengthened the category boundary as it got sharper.

On the other hand, the GR results showed influence of the mother tongue and the prior target language skills. There were within Day differences, but no training effects. First, the learners' native language /f/ is probably identical with the target language one, which is why they were rated similarly, irrespective of the foreign word context. The native language effect explains also the /f/ prototype and the non-prototypical training stimuli reflecting a hierarchy in the category. Again, the mother tongue effect can be seen in the GRs of the prototypical category representatives when the /f/ was rated better than the English /v/. Second, hierarchy in the baseline /v/ category probably indicates that the highly advanced learners have really learned to perceive the foreign language contrast. In addition, the fact that the prototypical /v/ still differed from the boundary stimulus in the final measurement, indicates also the same, that the subjects already master it. Third, the most interesting result is that training did not have any effects on the GR. This is probably due to the mother tongue influence on the /f/, and the existent hierarchy in the /v/, so both categories were hierarchical already in the baseline test.

Training effects were, however, seen in the decreased RTs and in the increased discrimination sensitivity. Naturally, the learners were able to discriminate the stimuli but the training also had an effect since the subjects got faster and more accurate in the discrimination of the voicing. Thus, training facilitates discrimination, even in the advanced learners.

Finally, prior target language skills and training effects can be seen in the MMN results. The baseline response showed that the advanced English university students perceived the difficult voicing contrast pre-attentively before training. It seems that the mother tongue influence was not strong at the pre-attentive level and that students perceived the contrast according to the target language. The pre-attentive perception of the voicing contrast changed, since the MMN response amplitude increased in the course of the listen-and-repeat training. The further strengthening of the memory traces was evident in the MMN response in the final measurement after four blocks of listen-and-repeat training. It resembled a prominent native-like response [8, 10, 12]. Which clearly demonstrates that a memory trace can be further enhanced. The latency of the response remained unchanged irrespective of the training. Perhaps the response was already as early as it could be.

To conclude, it is obvious that training has effects even on advanced language learners' perception. The behavioural improvement and the memory trace strengthening may indicate that the learning process is incomplete, as was the case in [8]. At the same time, this strengthening shows that not only the naïve learners [11, 5] can show training effects.

5. ACKNOWLEDGEMENTS

This study was financially supported by Emil Aaltonen Foundation. We wish to thank all the subjects who participated and M.A. Eetu Laine for his assistance with data collection.

6. REFERENCES

[1] Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., Näätänen, R. 1998. Development of language-specific phoneme representations in the infant brain. *Nat. Neurosci.* 1, 351–353.

[2] Flege, J. E., MacKay, I. R. A. 2004. Perceiving vowels in a second language. *Stud. Second Lang. Acquis.* 26, 1–34.

[3] Iverson, P., Pinet, M., Evans, B. G. 2012. Auditory training for experienced and inexperienced second-language learners: Native French speakers learning English vowels. *Appl. Psycholinguist.* 33, 145–160.

[4] Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., Lindblom, B. 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255, 606–608.

[5] Menning, H., Imaizumi, S., Zwitserlood, P., Pantev, C. 2002. Plasticity of the human auditory cortex induced by discrimination learning of non-native, mora-timed contrast of the Japanese language. *Learn. Mem.* 9, 253–267.

[6] Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., Alho, K. 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385, 432–434.

[7] Oldfield, R. C. 1971. The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychol.* 9, 97–113.

[8] Peltola, M. S., Kujala, T., Tuomainen, J., Ek, M., Aaltonen, O., Näätänen, R. 2003. Native and foreign vowel discrimination as indexed by the mismatch negativity (MMN) response. *Neurosci. Lett.* 352, 25–28.

[9] Peltola, M. S., Aaltonen, O. 2005. Long-term memory trace activation for vowels depends on the mother tongue and the linguistic context. *J. Psychophysiol.* 19, 159–164.

[10] Peltola, M. S., Tamminen, H., Toivonen, H., Kujala, T., Näätänen, R. 2012. Different kinds of bilinguals – Different kinds of brains: The neural organisation of two languages in one brain. *Brain Lang.* 121, 261–266.

[11] Tamminen, H., Peltola, M. S., Kujala, T., Näätänen, R. 2015. Phonetic training and non-native speech perception – new memory traces evolve in just three days as indexed by the mismatch negativity (MMN) and behavioural measures. *Int. J. Psychophysiol.* Accepted.

[12] Tamminen, H., Peltola, M. S., Toivonen, H., Kujala, T., Näätänen, R. 2013. Phonological processing differences in bilinguals and monolinguals. *Int. J. Psychophysiol.* 87, 8–12.

[13] Tremblay, K., Kraus, N., McGee, T. 1998. The time course of auditory perceptual learning: neurophysiological changes during speech-sound training. *NeuroReport* 9, 3557–3560.