

University of Tartu
Faculty of Social Sciences
Institute of Psychology

Anna-Magdalena Peterson

**Chinese native speakers' perception of Estonian quantity levels before and after
short-term training**

Research paper

Supervisors: Siqi Lyu

Running head: Language perception before and after short-term training

Tartu 2026

Chinese native speakers' perception of Estonian quantity levels before and after short-term training

Abstract

This study investigated whether short-term targeted language training can improve native Mandarin speakers' discrimination of Estonian quantity contrasts. 61 L1 Mandarin speakers were asked to complete a pre-test, a training phase, and a post-test where they had to discriminate Estonian quantity contrasts. The sample was divided into a control group and a training group. During the training phase, the training group was asked to discriminate Estonian quantity contrasts with feedback given, while the control group completed a visual perception task. The results proved that being in the training group gave an advantage in discrimination accuracy in the post-test, but no improvement was observed in the control group. The results implied that short-term targeted training can be effective in improving L1 Mandarin speakers' perception of Estonian quantity categories.

Keywords: short-term targeted training, Estonian quantity, Chinese lexical tone, AX discrimination task, language perception.

Hiina keelt emakeelena kõnelevate inimeste taju eesti keele vältevaheldusest enne ja pärast lühiajalist keeletreeningut

Kokkuvõte

Käesoleva töö eesmärk oli uurida, kas lühiajaline suunatud keeletreening võib parandada hiina keelt emakeelena kõnelevate võimet eristada eesti keele vältevaheldust. 61-l hiina emakeelega inimesel paluti läbida eeltest, treeningfaas ja järeltest, kus nad pidid eristama eesti keele välteid. Valim jagati kontrollrühmaks ja treeningrühmaks. Treeningfaasis paluti treeningrühmal teha eesti keele völdete eristamisel põhinevat tagasisidestatud harjutust, kontrollrühm aga tegi visuaalse taju ülesannet. Tulemused näitasid, et treeningrühmas olemine andis järeltestis eelise eristamise täpsuses, samas kui kontrollrühmas arengut ei täheldatud. Tulemused näitasid, et lühike suunatud treening võib olla tõhus hiina keelt emakeelena kõnelevate eesti keele vältevahelduste tajumise parandamisel.

Võtmesõnad: lühiajaline suunatud treening, eesti keele völted, hiina keele toon, AX diskrimineerimisülesanne, keeletaju.

Introduction

Human speech perception is strongly shaped by one's native language. That is why usually when speaking a foreign language, an accent appears: the processing of non-native contrasts is done within the structures of one's mother tongue. The more a language differs in its phonological structures from the native language, the harder it is to master. In this study, we examine the perception of Estonian quantity by Mandarin Chinese native speakers. The chosen languages are very diverse in phonological structures: Mandarin Chinese is based on lexical tones, whereas Estonian relies on phonemic quantity, using pitch only as a secondary cue (Lippus et al., 2009). We want to know if speakers of a pitch-dominant language can perceive subtle durational and tonal contrasts in a language like Estonian after only short-term targeted training.

Earlier studies suggest that early cortical processing of speech sounds is influenced by the salience of acoustic features in one's native language (Chandrasekaran et al., 2007). Kaan et al. (2007) demonstrated that targeted training can improve the sensitivity to new phonological contrasts. The importance of this study is in testing the effect of short-term targeted training on the ability of Mandarin speakers to overcome their perceptual biases and successfully discriminate between the short (Q1), long (Q2), and overlong (Q3) Estonian quantity categories. We hope this study will help in understanding cross-linguistic prosodic learning, and the plasticity of phonological perception systems in adult learners.

Mandarin Chinese Lexical Tone

Mandarin Chinese is a tonal language, which contains four lexical tones: high-level (ˉ), rising (ˊ), dipping or low (ˇ), and falling (ˋ) (Xu, 1997). These pitch contours distinguish word meanings — for example, *mā* (mother), *má* (hemp), *mǎ* (horse), and *mà* (scold). As sensitivity to pitch contour develops from an early age amongst native Mandarin Chinese speakers, it becomes deeply integrated into their phonological perception.

As being attuned to tonal cues gives Mandarin Chinese speakers an advantage in learning other pitch-oriented languages, the lack of familiarity with durational contrasts poses a challenge in mastering languages with multiple quantity features (He, 2010; Liu et al., 2022). Nevertheless, it has been shown that Chinese Mandarin speakers are able to use their pitch sensitivity in situations, where pitch is used differently than in their mother tongue (for example in languages where pitch contour is a secondary cue) (Laméris et al., 2024).

Estonian Quantity

Estonian is known for its unique three-way phonemic quantity contrast, distinguishing between short (Q1), long (Q2), and overlong (Q3) syllable durations. It means that the duration of a phonetic unit physically determines the quantity distinction. These quantity distinctions play a critical role in the lexical and grammatical structure of Estonian, often altering the meaning of the word (e.g. Q1 *koma*: [ko:ma] “comma”, Q2 *kooma*: [ko:.oma] “coma” and Q3 *kooma*: [ko;.oma] “to knit”). Quantity contrast does not solely apply to vowels but can also be noticed in syllable-medial consonants (e.g. *koli* [ko'li] “stuff”, *kolli* [kol'li] “monster’s”, *kolli* [kol':li] “monster”). Since in Estonian language, verbal stress is placed on the first syllable of the word, the phenomena of three-way quantity contrast are best shown on two-syllable words. As Lippus et. al. (2009) demonstrated in their study, overlong (Q3) syllables exhibit a falling pitch contour, which indicates that Estonian language involves a tonal component as a secondary cue. Therefore, Estonian quantity is best understood as a composite system, with duration being the primary, and pitch the secondary cue for phonological discrimination.

Research has shown that non-native listeners tend to rely more heavily on absolute duration, sometimes missing subtler tonal or relational cues crucial for accurate identification. (e.g., Leppik et al., 2018; Meister et al., 2015). However, it has not been studied too much on Chinese Mandarin speakers, who might show bigger sensitivity to pitch cues due to the significance of them in their native language.

Development of Language Perception and Foreign Language Learning

People’s language perception gets established early on in life. At the age of 6 to 12 months, children go through a sensitive period during which their perceptual system becomes attuned to the sounds of their native language (or languages, in multilingual environments), enabling them to distinguish phonemic contrasts relevant to their linguistic input while gradually losing sensitivity to non-native contrasts. Due to this process, known also as perceptual narrowing, the efficiency of processing native language patterns enhances, while the ability to distinguish unfamiliar phonemic contrasts degrades (Kuhl et al., 2008).

In adulthood, perceptual filters, that have been established in early childhood, can make learning a new language relatively challenging. Particularly due to the phonological features that are not used in the listener's first language. Although it deteriorates over time, studies show that even in adulthood neural plasticity is retained to some degree (Yang et al., 2015). Thanks

to that, new languages can be learned later in life but mastering their phonological features is much more difficult than in early childhood.

Foreign language learning models, like the Speech Learning Model (Flege, 1995) or the Perceptual Assimilation Model (Best & Tyler, 2007), suggest that new language contrasts are perceived through the lens of one's first languages' phonological system. This means that unfamiliar phonological phenomena are often assimilated into the closest native category.

However, if a language lacks certain types of phonological categories (e.g. Mandarin does not have durational cues), it is difficult to place the unfamiliar sound to any existing category. This is why learning a new language does not only need the skill to correctly acquire previously mastered sounds and language patterns but also overcome established perceptual boundaries.

Effect of Training on Foreign Language Learning

Although it might seem difficult to change something in the firmly established perceptual structures in the human brain, short-term targeted training has in fact been found to improve non-native phonological perception abilities, including distinguishing unfamiliar phonetic categories (Iverson et al., 2005). Short-term exposure, including feedback and variability, has been found to be effective in multiple studies about language perception and learning (Kaan et al., 2007; Li, 2016). However, the focus has been more on the perception of the quality of the phonetic units, (e.g. the ability of Japanese native speakers to learn English r/l distinction).

Quantity perception among non-quantity language speakers has been examined from some perspectives, although it is still very limited. A study by Ylinen et al. (2005) showed that the longer the exposure to durational cues, the better the level of understanding and discriminating them amongst foreign language learners. Nevertheless, even with a rather long exposure, foreign language learners could not master the durational cues completely. Another study showed that passive listening has been found to bring some effect upon discriminating quantity contrast (Kurkela et al., 2019). Recently, there have been studies that investigate how Chinese participants learn Estonian quantity (Lyu et al., 2025; Peelman, 2023). They adopted a passive training paradigm where participants listened to the training materials at home while reading a self-selected book. When looking at the overall perception sensitivity in the behavioral data, participants' perception improved in both the control and training groups

(Peelmann, 2023). Yet the training did not have a significant effect. It is suggested that passive training has a limited effect on participants' attentive perception of the Estonian quantity.

The Present Study

The present study is built on research by Lyu et al. (2025) and Peelmann (2023). Different from the previous studies where a passive training paradigm was applied, we will adopt an active training paradigm this time. We want to see whether a short-term targeted training period is sufficient to improve Mandarin speakers' perception of Estonian quantity distinctions.

H1: Chinese participants who receive the active training (i.e., the training group) will improve in discriminating the Estonian quantity.

H2: Chinese participants who do not undergo active training (i.e., the control group) will show no improvement in the pre- and post-test scores.

Method

Study Design

The participants completed two tests, first at the beginning of the session and second after a short-term auditory training of the Estonian quantity (training group) or a visual perception task (control group). Both the training group and the control group took the same pre-test and post-test. The two groups differed only in the training session in between. The pre-test, training, and post-test were conducted on the same day.

Participants

The sample consisted of 61 native Mandarin-speaking individuals. The participants were randomly assigned to two groups: a training group (N=31) and a control group (N=30) (see Table 1). All of the participants were right-handed and did not report hearing impairment problems.

Table 1*Demographic Information of the Participants in Both Groups*

	Females	Males	Mean (age)	SD (age)	R (age)
Control group	15	15	22.93	2.03	20 - 27
Training group	15	16	22.45	1.82	19 - 25

Note. *SD* = *Standard Deviation*; *R* = *Range*.

Stimuli

The stimuli were disyllabic Estonian pseudowords. The stimuli were first created by constructing different consonant-vowel combinations. The first syllable was combined with consonants (h, j, k, l, m, n, p, r, s, t, v) and vowels (a, e, i, o, u, ä, õ, ö, ü), which gave 99 syllable combinations. The second syllable was constructed by consonants (d, g) and vowels (a, e, i, u), which gave 8 combinations. Combining all the syllables with each other made a total of 792 disyllabic base words. Each base word was put into three quantity degrees (Q1, Q2, Q3). The list of words underwent filtering to eliminate words with actual meanings. Firstly, the list was matched to *sõnaveeb.ee* wordlist, which has 7 million Estonian words. Secondly, 3 adult native Estonians proofread the remaining list to eliminate the words that could not be found on the *sõnaveeb.ee* list. Initially, 298 sets of words were recorded for the stimuli, from which 60 sets will be chosen for the experiment. Both pre- and post-tests and the training were conducted with the same stimuli.

Each set of stimuli consisted of the variants of Q1, Q2, and Q3. Additionally, a manipulated flattened Q3 (FQ3) was created, resulting in four variants of stimuli in each set. The FQ3 was constructed by extracting the pitch contour of Q2 and applying it to Q3, so that it carries the Q3 duration but a Q2 (flat) pitch. The inclusion of FQ3 is for the purpose of another project and will not be analyzed in this paper.

Procedure

The experiment was conducted in a quiet and dimly lit laboratory. Participants were given headphones and a computer, which was utilized for receiving information and giving answers. Before the experiment, participants received instructions in Chinese orally and on the computer screen.

The pre- and post-tests adopted an AX discrimination task which is a commonly used design in speech research, also known as the same - different (AX) design (Gerrits & Schouten, 2004). This type of task has two stimuli in each trial, in this case pseudowords, paired together. The words in the pair might be different or the same, and the participant must distinguish whether the two sounds are the same or different. A trial started with a cross “+” on the screen that lasted for 300 ms. After that, participants heard the first stimulus (AUDIO A), followed by an interstimulus interval (ISI) of 575 ms and then the second stimulus (AUDIO X). Participants were instructed to respond as fast as possible after hearing the second stimulus. If the participant did not respond within 2000 ms, the program proceeded to the next trial. Participants’ reaction time and accuracy were recorded.

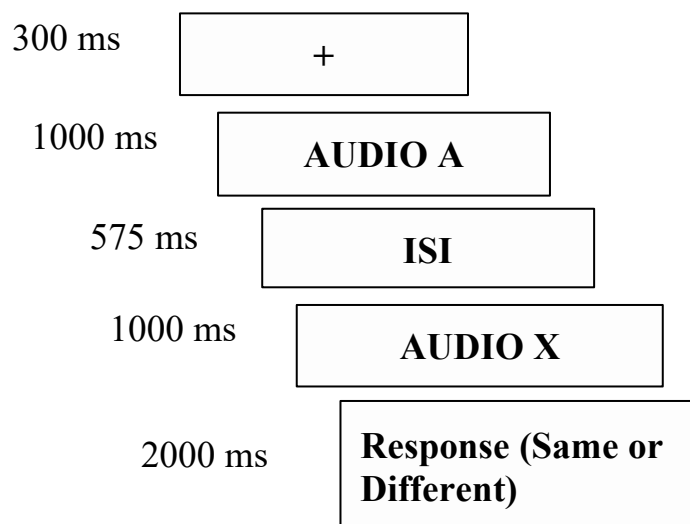


Figure 1. AX discrimination task trial procedure

The AX experiment consisted of both different pairs and same pairs. For each set of stimuli, we created four different pairs and four of the same pairs, as shown in Table 2.

Table 2*Different and same word pair combinations*

Pair type	Combination 1	Combination 2	Combination 3	Combination 4
Different	Q1 vs. Q2	Q2 vs. Q3	Q3 vs. FQ3	Q1 vs. FQ3
Same	Q1 vs. Q1	Q2 vs. Q2	Q3 vs. Q3	FQ3 vs. FQ3

Regarding the training phase, the training group completed an adaptive training task on the same stimuli that was used in the pre- and post-tests. Participants were asked to accurately respond to the stimulus pairs in order to eliminate them from the experiment. If the participant failed to do so, the stimuli stayed in the test until the participant got it right. The control group participated in a visual perception task.

Results

Two 2×2 mixed-design ANOVA analyses were conducted using JASP-0.95.4.0 software to examine whether short-term targeted training improved Chinese participants' discrimination of Estonian quantity contrasts in terms of accuracy and reaction time. In both models, Time (pre-test, post-test) was treated as a within-subjects factor, and Group as a between-subjects factor.

The first ANOVA test was conducted to analyze the effects of Group and Time on Accuracy. The mean and SD for accuracy for each group are given in Table 3, and the statistical output is given in Table 4. There was no significant main effect of Time, indicating that overall accuracy did not significantly differ between the pre- and post-test when collapsing across groups. There was, however, a significant main effect of Group, with the training group showing higher overall accuracy than the control group in both the pre- and post-tests (see Table 4).

The analysis showed a significant Time $\times$ Group interaction, meaning that the change in accuracy from pre-test to post-test differed between the groups (see Figure 2). Post-hoc tests with Holm adjustment were conducted to explore the interaction between Group and Time (pre- or post-test) on accuracy. Within group comparisons revealed a significant decrease in

accuracy from pre-test to post-test in the control group, $p < .001$, whereas the change observed in the training group was marginal, ($p = .077$) and did not reach statistical significance. The training group showed significantly higher accuracy than the control group at post-test.

Table 3

Means and Standard Deviations for Accuracy by Group and Time

Group	Time	N	Mean Accuracy	SD
Control	Pre-test	30	0.880	0.077
	Post-test	30	0.842	0.098
Training	Pre-test	31	0.900	0.078
	Post-test	31	0.920	0.054

Table 4

Summarized Results of Two-Way ANOVA Tests for Accuracy

Group	SS	df	MS	F	p
Within-subject effects					
Time	0.003	1	0.003	2.33	.132
Time x Group	0.025	1	0.025	22.46	< .001
Residuals	0.065	59	0.001		
Between-subject effects					
Group	0.073	1	0.073	6.61	.013
Residuals	0.654	59	0.011		

Note. SS = Sum of Squares; MS = Mean Square. Time = pre-test vs. post-test accuracy.

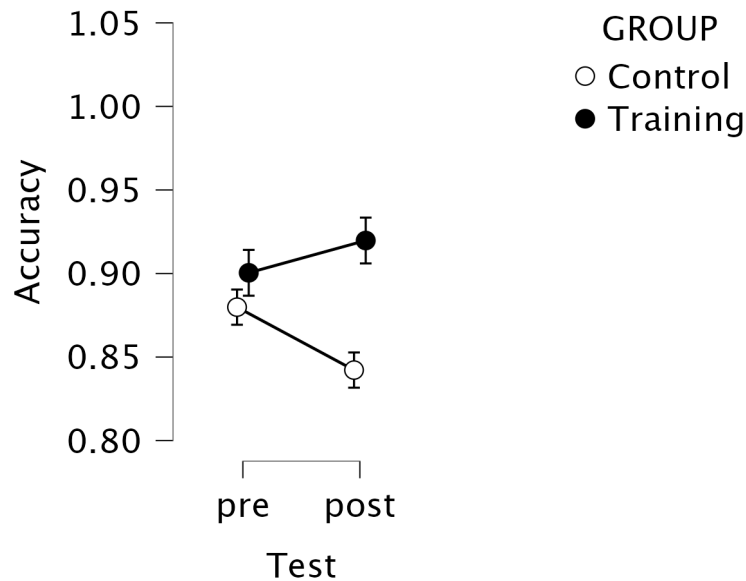


Figure 2. Accuracy x Group interaction in pre- and post-tests

Note. Confidence intervals 95%.

The second ANOVA test was conducted to analyze the effects of Group and Time on Reaction Time. The mean and SD for reaction time for each group are given in Table 5, and the statistical output is given in Table 6. There was a significant main effect of Time, indicating that reaction times decreased from pre-test to post-test across both groups (see Table 6). The main effect of Group was also significant, with the training group showing overall slower reaction times than the control group in both the pre- and post-tests. The Time $\times$ Group interaction was not significant, indicating that the magnitude of pre-post reaction time change did not differ between the training and control groups. Both groups showed similar reductions in response latency over time (see Figure 3).

Table 5*Means and Standard Deviations for Reaction Time (Seconds) by Group and Time*

Group	Time	N	Mean	SD
Control	Pre-test	30	0.820	0.093
	Post-test	30	0.748	0.086
Training	Pre-test	31	0.886	0.120
	Post-test	31	0.827	0.109

Table 6*Summarized Results of Two-Way ANOVA Tests for Reaction Time*

Group	SS	df	MS	F	p
Within-subject effects					
Time	0.129	1	0.129	62.327	< .001
Time x Group	0.001	1	0.001	0.552	.460
Residuals	0.123	59	0.002		
Between-subject effects					
Group	0.158	1	0.158	8.303	.006
Residuals	1.126	59	0.019		

Note. SS = Sum of Squares; MS = Mean Square. Time = pre-test vs. post-test accuracy.

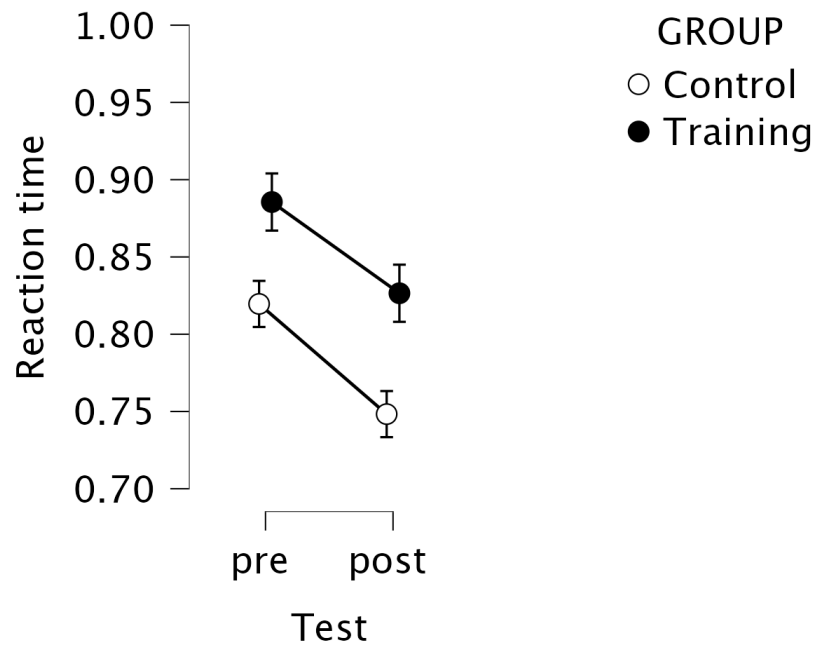


Figure 3. Control and training group reaction time across pre- and post-tests in seconds.

Note. Confidence intervals 95%

Discussion

The purpose of this study was to determine whether short-term targeted training can improve the discriminating of Estonian quantity contrasts by Mandarin Chinese native speakers. The following hypotheses were tested: (H1) Chinese participants who receive the active training (i.e., the training group) would improve in discriminating the Estonian quantity; (H2) Chinese participants who do not undergo active training (i.e., the control group) will show no improvement in the post-test scores. We found that the active training improved participants' accuracy in discriminating Estonian quantity contrasts, whereas no improvement was observed in the control group. The results support our hypotheses.

Active Training Effect on Native Chinese Speakers in Discriminating Estonian Quantity Contrasts

Our results revealed an advantage for participants who received training, but no such improvement was found in the control group, which did not receive any training. Our findings suggest that short-term training is effective in learning Estonian quantity.

This finding is in accord with earlier findings about short-term targeted training (e.g. Kaan et al., 2007 and Li, 2016), where it was found that short-term training paradigms can help in many areas of distinguishing unfamiliar phonemic units. A meta-analysis by Yao and colleagues (2025) revealed a large overall positive effect of phonetic training on L2 phonetic competence across diverse studies and instructional contexts, finding that perceptual training had a larger effect size compared to production or combined training. The current study helped to broaden the line of perceptual training research by examining improvement in the perception of three-way quantity contrasts by tonal language L1 speakers, an area of research which has received limited attention. The training period analyzed in this study was also unique in its duration, as many previous studies in language perception have implemented training over multiple days. This study's experiment sessions were completed within a single day, and the participants showed measurable improvement already after only one training session. In the light of previous research, it might be suggested, that in very little time, highly targeted training may be sufficient to prompt short-term perceptual gains in the perception of complex quantity contrasts, although longer training has probably stronger effects.

It is important to note that the training group's improvement was marginal, improving from a mean accuracy of 0.90 in the pre-test to 0.92 in the post-test. We suggest that the

marginal improvement could be the consequence of the Chinese participants' already high performance at pre-test, which left limited room for further improvement. This ceiling effect, a methodological complication that is often encountered in speech and language perception research (Bruijnzeel et al., 2017; Iverson, 2005; Kaan, 2007) may have contributed to the reduced statistical strength of the observed training effect.

The unexpectedly high pre-test accuracy may be attributable to properties of the stimulus set, particularly the inclusion of the overlong Q3 contrast, which contained a salient falling pitch cue. Pitch contour is a primary phonological cue in Mandarin Chinese and plays a critical role in lexical meaning (Liu et al., 2022). The Q3 stimulus may have been easier for L1 Mandarin speakers to perceive than contrasts differing solely in duration. In consequence, the inclusion of Q3 may have facilitated discriminating between different quantity contrasts and limited observable learning effects. Future studies aiming to isolate purely durational perception in L1 Mandarin speakers may therefore benefit from excluding stimuli that contain pitch cues.

An additional interesting finding was the control group's significant decrease in accuracy in the post-test, suggesting that without reinforcement or exposure, discrimination of Estonian quantity contrasts may be unstable for native Mandarin listeners. This pattern aligns with research showing that perceptual sensitivity to non-native contrasts can decline in the absence of training or continued engagement, particularly for phonetic contrasts that are not salient in the listener's L1 phonological system (e.g., Iverson et al., 2005). Another explanation for the decrease in accuracy might be that the visual perception task was cognitively demanding and that participants were fatigued. Sustained attention studies have shown that performance on perceptually demanding discrimination tasks can decrease over time due to vigilance decrement, pointing to cognitive fatigue, which leads to reduced accuracy over time (Pattyn et al., 2008).

Active and Passive Training on Foreign Language Learning

The findings of this study give valuable insight into the difference in quality between passive and active learning paradigms. Peelman (2023) demonstrated in a similarly designed study the effect of passive training of discriminating Estonian quantities by Mandarin - Chinese L1 speakers, which did not prove to be efficacious. The participants received passive language training for five days, in which they were asked to listen daily to Estonian quantity contrast while reading their favorite book. In contrast to this study, the tendency for both groups was to

improve in the discrimination task; it did not matter whether the participant was assigned to the control or the training group. The gains of the training group were not statistically significant though, and thus passive training proved to be ineffective. Differently, in the current study, although the training was only conducted on one day, the active training paradigm applied in this study proved to be effective in just a short period of time, the difference between the control and training group's accuracy in the post-test was statistically significant (Lyu et al., 2025; Peelman, 2023).

Although the findings on the effectiveness of passive language exposure are varied, there are studies which show that passive training can facilitate auditory perception of foreign linguistic features (Liu et al., 2025). The effects of passive and active auditory training have been studied with the help of mismatch negativity (MMN), which is a pre-attentive event-related potential that reflects the brain's automatic detection of a deviation from an established auditory regularity (Näätänen et al., 2007). A study by Kurkela et al. (2019) demonstrated that passive exposure to Mandarin lexical tones shortened the latency of MMN of Finnish listeners over a short period of passive training, which hints to a change in perceptual abilities.

Given these findings, it may be argued that the effectiveness of foreign language perceptual learning is influenced not only by the presence of training but also by the type of training employed. Wright et al. (2015) examined the effects of passive exposure, active training, and their combination. The findings showed greater perceptual gains for participants who received both active practice and passive exposure than for those who engaged in active training alone, while only passive exposure did not result in significant improvement. In the light of the present study, these findings are informative in suggesting that although active training has a significant and irreplaceable role in acquiring new phonemic units, additional passive exposure may contribute to more robust and sustained perceptual learning of a new linguistic system.

Task Familiarity Effects on Performance in an AX discrimination Task

Both the control and training groups in the present study showed reduced reaction times from pre-test to post-test, which we suggest could be due to task familiarity. Task familiarity has two components – familiarity with task content and task type – both of which have been shown to significantly improve performance (Révész & Han, 2009). In this study, pre- and post-tests used the same stimuli (i.e., task content), and the training group also rehearsed with these stimuli. Familiarity with the task content may have contributed to faster responses even

without active training. Additionally, the task type was identical across pre- and post-tests. Although instructions and required actions were simple (e.g., press “F” if the words differ, “J” if they are the same), participants may have needed time to adjust to key locations, response timing, and the experimental setting. The novelty of the pre-test likely slowed reactions, while familiarity in the post-test contributed to faster responses.

Our results suggest that reaction time may not be the most reliable measure of perceptual performance in AX discrimination tasks. In same–different tasks, reaction times are affected not only by perceptual sensitivity but also by decision strategies, response biases, and processing demands unrelated to the acoustic contrast (Pisoni & Tash, 2012). AX trials require holding the first stimulus in short-term memory, making reaction times sensitive to memory load, attention, and fatigue rather than purely perceptual differences (Gerrits & Schouten, 2004).

Limitations and Further Implications for Foreign Language Learning

It must be noted that many of the topics discussed in this paper could be studied more carefully. This study was conducted with a relatively small sample, which may limit the representativeness of the findings, especially with such fine-grained numerical changes as in accuracy or reaction time.

In addition, the latency of the effect needs further investigation. The present study only tested the endurance of the training effect for one day, however previous research suggests that it is likely that the impact of the training would diminish over time (Saito & Chen, 2025). Future studies should therefore examine the latency and retention of training effects over longer periods, for example by using delayed post-tests to determine how long improvements persist. It would also be valuable to investigate the amount and distribution of training required to achieve stable, long-term improvement in the perception of unfamiliar quantity contrasts.

Despite these limitations, the present study contributes to the growing body of research on short-term targeted phonemic training by examining Estonian quantity perception, a relatively underexplored domain in perceptual learning research. As such, it provides a foundation for future work on the acquisition of complex prosodic and quantity-based contrasts in foreign language learning.

Acknowledgements

I would like to express my sincere gratitude towards my supervisor Siqu Lyu, without whom this paper would not exist. Her patience, expertise, and thoughtful guidance supported me at every stage of the research process. I am truly grateful for the time and energy she dedicated to introducing the topic of language research, discussing ideas, reviewing drafts, and helping me find my way throughout the process.

References

- Best, C. T., Tyler, M., Bohn, O., & Munro, M. (2007). Nonnative and second-language speech perception. *Language experience in second language speech learning*, 17, 13-34. <https://doi.org/10.1075/llt.17.07bes>
- Bruijnzeel, H., Cattani, G., Stegeman, I., Topsakal, V., & Grolman, W. (2017). Incorporating ceiling effects during analysis of speech perception data from a paediatric cochlear implant cohort. *International Journal of Audiology*, 56(8), 550–558. <https://doi.org/10.1080/14992027.2017.1311029>
- Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007). Mismatch negativity to pitch contours is influenced by language experience. *Brain Research*, 1128(1), 148–156. <https://doi.org/10.1016/j.brainres.2006.10.064>
- Flege, J. E. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.
- Gerrits, E., & Schouten, M. E. H. (2004). Categorical perception depends on the Discrimination task. *Perception & Psychophysics*, 66(3), 363–376. <https://doi.org/10.3758/BF03194885>
- He, L. (2010). Perception and production of isolated and coarticulated Mandarin tones by American learners (Doctoral dissertation, University of Illinois at Urbana-Champaign).
- Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *The Journal of the Acoustical Society of America*, 118(5), 3267–3278. <https://doi.org/10.1121/1.2062307>
- Kaan, E., Wayland, R., Bao, M., & Barkley, C. M. (2007). Effects of native language and training on lexical tone perception: An event-related potential study. *Brain Research*, 1148, 113–122. <https://doi.org/10.1016/j.brainres.2007.02.019>
- Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2008). Early speech Perception and later language development: Implications for the “critical period.” *Language Learning and Development*, 1(3-4), 237–264. DOI:10.1207/s15473341lld0103&4_2
- Kurkela, J. L., Hämäläinen, J. A., Leppänen, P. H., Shu, H., & Astikainen, P. (2019). Passive

- exposure to speech sounds modifies change detection brain responses in adults. *NeuroImage*, 188, 208-216. doi: 10.1016/j.neuroimage.2018.12.010
- Laméris, T. J., Llompart, M., & Post, B. (2024). Non-native tone categorization and word learning across a spectrum of L1 tonal statuses. *Bilingualism: Language and Cognition*, 27(4), 729-743. DOI: <https://doi.org/10.1017/S1366728923000871>
- Leppik, M., Lippus, P., & Asu, E. L. (2018). The perception of Estonian quantity degrees By Spanish listeners. *Conference on Speech Prosody* (pp. 478-482)..
- Li, Y. (2016). Effects of high variability phonetic training on monosyllabic and disyllabic Mandarin Chinese tones for L2 Chinese learners (Doctoral dissertation, The Ohio State University).
- Lippus, P., Asu, E. L., & Allik, J. (2009). The tonal component of Estonian quantity in native and non-native perception. *Journal of Phonetics*, 37(4), 388–396. <https://doi.org/10.1016/j.wocn.2009.07.002>
- Liu, L., Lai, R., Singh, L., Kalashnikova, M., Wong, P. C., Kasisopa, B., ... & Burnham, D. (2022). The tone atlas of perceptual discriminability and perceptual distance: Four tone languages and five language groups. *Brain and Language*, 229, 105106. <https://doi.org/10.1016/j.bandl.2022.105106>
- Lyu, S., Põldver, N., & Kreegipuu, K. (2025). Exploring passive training and training generalization in the learning of Chinese tone and Estonian quantity [Manuscript in preparation]. Institute of Psychology, University of Tartu.
- Meister, E., Nemoto, R., & Meister, L. (2015). Production of Estonian quantity contrasts by Japanese speakers. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 6(3), 79-96. DOI: <https://doi.org/10.12697/jeful.2015.6.3.03>
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, 118(12), 2544–2590. <https://doi.org/10.1016/j.clinph.2007.04.026>
- Saito, K., & Chen, S. (2025). Optimizing L2 phonetic learning: Spaced vs. massed training schedule. *Studies in Second Language Acquisition*, 1–22. <https://doi.org/10.1017/S02722263125101447>
- Pattyn, N., Neyt, X., Henderickx, D., & Soetens, E. (2008). Psychophysiological investigation of vigilance decrement: Boredom or cognitive fatigue? *Physiology & Behavior*, 93(1–2), 369–378. doi: 10.1016/j.physbeh.2007.09.016
- Peelmann, H. (2023). The training effect on the perception of Estonian quantity by Chinese

- native speakers. [Unpublished research paper]. University of Tartu.
- Pisoni, D. B., & Tash, J. (1974). Reaction times to comparisons within and across phonetic categories. *Perception & Psychophysics*, 15(2), 285–290.
<https://doi.org/10.3758/BF03213946>
- Révész, A. Han ZH., (2006) Task Content Familiarity, Task Type and Efficacy of Recasts. *Language Awareness*, 15:3, 160-179, DOI: 10.2167/la401.0
- Wright, B. A., Baese-Berk, M. M., Marrone, N., & Bradlow, A. R. (2015). Enhancing speech learning by combining task practice with periods of stimulus exposure without practice. *The Journal of the Acoustical Society of America*, 138(2), 928–937.
<https://doi.org/10.1121/1.4927411>
- Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25(1), 61–83.
<https://doi.org/10.1006/jpho.1996.0034>
- Yang, J., Gates, K. M., Molenaar, P. C. M., & Li, P. (2015). Neural changes underlying successful second language word learning: An fMRI study. *Journal of Neurolinguistics*, 33, 29–49. <https://doi.org/10.1016/j.jneuroling.2014.09.004>
- Yao, Y., He, M., Chen, F., & Zhu, J. (2025). A meta-analysis of second language phonetic training: Exploring overall effect and moderating factors. *Journal of Speech, Language, and Hearing Research*, 68(4), 1784–1802.
https://doi.org/10.1044/2024_JSLHR-24-00432
- Ylinen, S., Shestakova, A., Alku, P., & Huotilainen, M. (2005). The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of Finnish. *Language and Speech*, 48(3), 313–338.
<https://doi.org/10.1177/00238309050480030401>

Käesolevaga kinnitan, et olen korrektselt viidanud kõigile oma töös kasutatud teiste autorite poolt loodud kirjalikele töödele, lausetele, mõtetele, ideedele või andmetele.

Olen nõus oma töö avaldamisega Tartu Ülikooli digitaalarhiivis.

Anna-Magdalena Peterson