L2 Development of Quantity Perception: Dutch Listeners Learning Finnish /t-tː/

Willemijn Heeren

Utrecht institute of Linguistics OTS
Utrecht University, Utrecht, The Netherlands
Willemijn.Heeren@let.uu.nl

10.21437/Interspeech.2005-350

Abstract

The perceptual development of Dutch listeners learning to perceive the Finnish quantity contrast /t-tː/ was studied. It is shown that short laboratory training is (i) sufficient to change identification of relevant speech sounds, but (ii) insufficient to substantially change perceptual sensitivity along the phoneme continuum. Furthermore, L2 learners need much more relevant language experience, but can obtain native-like sensitivity to the /t-tː/ contrast.

1. Introduction

In quantity languages, segment length is phonologically contrastive, and is acoustically cued by segment duration. Throughout the world's languages, vowel length contrasts are found more often than consonant length contrasts [1]. Consonant length contrasts occur, for instance, in Italian, and a language such as Finnish or Japanese has both contrastive vowel and consonant length.

Dutch is not a quantity language, but it does have segments that contrast in length. The language has short-long vowel pairs that are, however, also cued by spectral differences between the vowels. The long vowels are 1.5 to 2 times longer in duration than the short ones [2, 3]. Dutch has no short-long contrast for consonants. But consonant duration can be influenced by the length of a preceding vowel [4, 5, but see also 6]. This difference, however, was in the range of only 5 to 12 ms, whereas double consonants in quantity languages are about twice as long as singletons. Consonant duration in Dutch can further be influenced by surrounding consonants [7]: the more complex the onset or coda, the shorter the duration of the consonants constituting it.

Thus, even though Dutch does employ some durational differences on the segmental level, it does not use consonant length contrastively in the way quantity languages do. Accordingly, it is expected that Dutch listeners will have difficulty discriminating pairs of words differing in consonant duration alone, at least in comparison with Finnish listeners, whose native language uses consonant length contrastively.

The present study addressed the question of how Dutch adult perception of the Finnish quantity contrast /t-tː/ develops as a function of training. We hypothesize that learning occurs through Acquired Distinctiveness or through Acquired Similarity [8]. In case of acquired distinctiveness (AD), the nonnative listener initially has difficulty discriminating the nonnative phonemes, but learns to distinguish speech sounds that he is trained to categorize differently. In case of acquired similarity (AS), the nonnative listener initially discriminates between-category and within-category speech sounds. As a result of training, however, only speech sounds that he is trained to categorize differently remain easy to discriminate. In other words, the main question is whether the perception of a nonnative phoneme contrast develops through an increase in sensitivity at the newly learned phoneme boundary, or through a decrease in sensitivity within the newly learned phoneme categories.

2. Experiment I

In a training study, Dutch listeners learned to classify the nonnative speech sounds /t/ and /tː/ into two phoneme categories similarly to native Finnish listeners. A pretest-training-posttest design was used. Before training, baseline perception was measured in a classification and a discrimination test. After training, these same tests were run again. Furthermore, control data were collected from Dutch listeners who did not learn the Finnish phoneme contrast. It was expected that trained listeners learn through acquired distinctiveness, which is supported by earlier work within this research project [9].

2.1. Method

2.1.1. *Materials*

The Finnish pseudoword pair *ata-atta* (/ɑtɑ-ɑtːɑ/) was recorded spoken by six different, male and female, speakers. Six phoneme continua of 7 stimuli each were made by shortening the closure duration of /tː/ to /t/ in equal steps. One female speaker's continuum was used as test continuum and the other speakers' continua were used as training materials.

2.1.2. *Participants*

A group of 28 students from Utrecht University participated, all native speakers of Dutch without hearing problems. Their mean age was 23 years. Half of them formed the test group, completing the entire procedure – pretest-training-posttest – whereas the other half were the controls, doing only pretest and posttest without receiving any training. Furthermore, 6 Finnish listeners provided control data.

2.1.3. *Design*

In pretest and posttest, listeners only heard the test speaker: firstly they completed classification of the /t-tː/ continuum, followed by 2IFC discrimination at 1- and 2-step stimulus distances. During this task, participants respond by giving the presentation order of two different stimuli taken from the continuum, so either AB or BA. Two different step sizes were used to give both hypotheses, AD and AS, a chance. If discrimination performance improves as a function of training (reflected by an increase in the percentage of correct responses), this is may show up mainly in 1-step discrimination, since 1-step differences are generally difficult to discriminate. On the other hand, if discrimination performance falls as a result of training, this may mainly show in reduced percentages of correct responses to 2-step differences that are generally relatively easy to discriminate.

As training, classification sessions with feedback were run, using phoneme continua based on speech from 5 speakers.

2.1.4. Procedure

Tests were run in a quiet room. A laptop computer was used both to present the stimuli in a random order and to register responses. Stimuli were presented over headphones at a comfortable listening level. Listeners were tested individually and received written instructions.

In the pretest, listeners classified the continuum by labeling each stimulus as containing a ‘short t’ or a ‘long t’. During a short introduction the continuum’s endpoints were each presented twice. The test itself consisted of 10 stimulus repetitions, i.e. 70 trials, and included a break halfway. This lasted about 4 minutes. Next, 2IFC discrimination with 1-step and 2-step stimulus pairs was administered. The inter-stimulus interval was 200 ms. Listeners responded by indicating which presentation order they had heard: ‘the first word contained the shorter [t] and the second the longer [t]’ or the other way around. As an introduction, a few examples of 6-step stimulus pairs were presented, followed by 11 practice trials. The actual test consisted of 20 repetitions per stimulus pair (10 in each order). Stimulus order was randomized and three breaks were included at regular intervals. The test lasted for 15-20 minutes.

Between pretest and posttest, listeners received training until they reached a score of 90% correct responses in two subsequent 350-trial training sessions, each containing 10 repetitions of 7-stimulus continua from 5 different speakers. The posttest was of exactly the same content as the pretest, and was administered one day after the last training.

2.2. Analysis & results

Pretest and posttest classification results are represented as percentages of /tː/ responses to each of the duration steps per listener. These percentages were transformed to z-scores, which were subjected to linear regression analyses, per listener and per test. From the regression fits, the phoneme boundary and boundary width were derived. The phoneme boundary was fixed at the 50% point where $z = 0$. The boundary width was determined by subtracting the 25% point (at $z = -.674$) from the 75% point (at $z = .674$).

2IFC discrimination results are given in percentages of correct responses to each of the stimulus pairs.

The mean amount of training necessary for proceeding to the posttest was 1625 trials, ranging from 700 to 3150 trials.

2.2.1. Classification of the test continuum

The phoneme boundaries and boundary widths were subjected to repeated measures analyses with within-subjects factor Test (pretest vs. posttest) and between-subjects factor Listener Group (trained vs. control). For reference: Finnish listeners put the boundary at a closure duration of 203 ms with a boundary width of 35 ms.

The phoneme boundaries only showed a main effect of Listener Group [$F(1,26) = 7.2$, $p = .013$]. Trained listeners’ phoneme boundaries shifted from pretest to posttest (223 to 215 ms). The control listeners’ boundary remained at the same location (233 ms in pretest and posttest). The pretest difference between the listener groups was not significant, but the posttest difference between the listener groups was significant. The trained listeners’ boundary shifted towards that of the Finnish as a result of training.

Furthermore, the boundary width of the trained listeners decreased as a result of training (from 39 to 33 ms), whereas this was not the case for the controls (38 ms in both tests). The boundary widths showed a main effect of Test and a Test × Listener Group interaction [$F(1,26) = 5.2$, $p = .031$; $F(1,26) = 4.9$, $p = .035$]. The trained group’s boundary width was reduced by training, while the control group’s boundary width remained the same.

2.2.2. 2IFC Discrimination of the test continuum

The discrimination test was included to monitor the changes in perceptual sensitivity along the test continuum as a result of training. The results of 1-step and 2-step discrimination were each subjected to a repeated measures ANOVA with Test (2) and Stimulus Pair (1-step: 1-6, or 2-step: 1-5) as within-subjects factors and Listener Group (trained vs. control) as between-subjects factor. If necessary, degrees of freedom were Huynh-Feldt corrected.

In 1-step 2IFC discrimination, a main effect of Stimulus Pair and a Test × Stimulus Pair interaction were found, [$F(4.3,112.2) = 21.3$, $p < .001$; $F(5,130) = 4.4$, $p = .001$]. Overall, trained listeners’ mean percentages correct *increased* somewhat from pretest to posttest (69% to 71%), whereas results for control listeners showed a small *decrease* (70% to 69%). But no effect of Listener Group or Test × Listener Group interaction was found, meaning that trained listeners’ perceptual sensitivity did not change compared to the controls’ sensitivity. Pretest-posttest results per listener group are shown in Figure 1. Data from six Finnish listeners are shown as a reference line.

It was expected that the trained listeners would develop heightened perceptual sensitivity around the location of their phoneme boundary, i.e. near stimulus pair 3-4 (201-236 ms). This peak is not present, but trained listeners did tend to give more correct answers to stimulus pairs 2-3 to 4-5 after training, i.e. the region around the phoneme boundary. This may mean that their perceptual sensitivity shifted to the relevant region, but the change was not significant. The control listeners’ discrimination results show a peak at stimulus pair 2-3 (166-201 ms) that lies significantly higher than their pretest responses to that stimulus pair. This peak, however, does not occur in the region where a higher perceptual sensitivity would be expected, i.e. near their perceived boundary at 233 ms.

As for 2-step discrimination, main effects of Test and Stimulus Pair were found [$F(1,26) = 9.8$, $p = .004$; $F(3.5, 90.8) = 42.4$, $p < .001$]. Post-hoc analyses revealed that only trained listeners’ discrimination levels increased significantly.

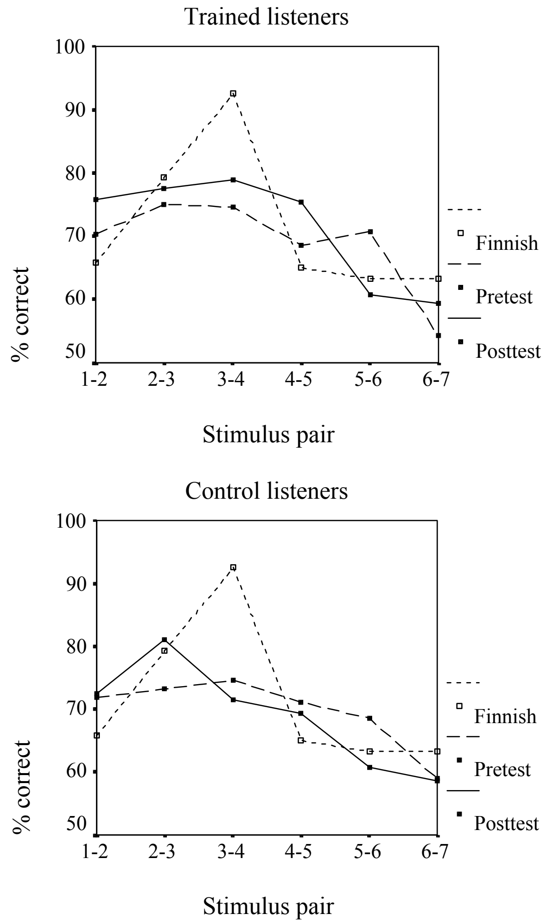

Figure 1: 1-step 2IFC discrimination results for trained listeners (upper) and controls (lower) with Finnish discrimination results as a dashed reference line in both.

2.3. Discussion

The present experiment was designed to answer the question how Dutch adult perception develops while learning the Finnish quantity contrast /t-tː/? It was expected that Dutch listeners would learn to perceive the new phoneme contrast through acquired distinctiveness, and not through acquired similarity.

The classification results showed that training with multiple speakers was effective, as had been found earlier [10], since trained listeners partially transferred the newly learned phoneme boundary from the training speakers to the test speaker. Trained listeners' phoneme boundaries moved closer to that of the Finnish listeners, whereas for controls this was not the case. Furthermore, only the trained speakers' boundary width decreased, reflecting their increased classification consistency.

Classification training had improved nonnative listeners' classification of the test continuum, but 1-step discrimination results did not resemble the native pattern: increased perceptual sensitivity near the phoneme boundary was not found for the language learners, who only *tended* to increase sensitivity in the relevant region. In 2-step discrimination, however, percentages of correct responses increased significantly for the trained listeners, but not for the controls. Since Dutch listeners' discrimination scores in the 1-step test were relatively low before training and tended to increase through training, and since no decrease in 2-step discrimination levels was found, development most likely occurs through acquired distinctiveness. This assumption is also supported by findings from earlier training studies [9, 11].

In sum, the findings from experiment I suggest that Dutch learners of Finnish need relatively little training to improve identification of ambiguous words varying in the closure duration of [t], but need much more experience to achieve native-like sensitivity along the /t-tː/ continuum. This last hypothesis was tested in experiment II.

3. Experiment II

If more experience is important to obtain native-like perceptual sensitivity along a quantity continuum, we expect Dutch learners of Finnish to increasingly show Finnish-like behavior. The question this experiment addresses is: what is the influence of experience on the perception of a Finnish quantity contrast by Dutch learners of Finnish? It is expected that advanced learners will increasingly show Finnish discrimination behavior, and that their development will occur through acquired distinctiveness.

3.1. Method

3.1.1. Participants

Eight students of Finnish (3 first-year, 3 second-year and 2 third-year students) from the University of Groningen, with Dutch as their native language, participated. Their mean age was 22. Six Finnish control listeners, living or working in the Netherlands at the time of testing, were also tested. Their mean age was 42 years.

3.1.2. Procedure

The listening tests with the students were run in a quiet room at the University of Groningen. Finnish listeners were tested at their homes or at work. The procedure was the same as in the pretest from experiment I (see section 2.1.4). Classification and 2IFC discrimination were completed within 45 minutes.

3.2. Analysis & results

The results were analyzed as in section 2.2. The behavior of students at three levels of learning was directly compared to that of the Finnish control listeners.

3.2.1. Classification of the test continuum

The effect of language experience on perception of the *ata-atta* continuum was addressed in an ANOVA with Closure Duration (131-341 ms) and Amount of Exposure (1^{st}, 2^{nd}, 3^{rd} year students, Finnish natives) as between-subjects variables. Firstly, a Closure Duration × Amount of Exposure interaction was found [$F(18,70) = 3.4$, $p < .001$]. Further, a main effect of Amount of Exposure and an effect of Closure Duration on the mean percentages of /tː/ responses were found [$F(3,70) = 11.2$, $p < .001$; $F(6,70) = 206.3$, $p < .001$]. Post-hoc tests showed that the third-year students' classification functions did not significantly differ from those of the native Finnish, whereas both other student groups did differ from the Finnish classification function. Apparently, advanced students of Finnish learn to classify the continuum similarly to native listeners.

3.2.2. 2IFC discrimination of the test continuum

Figure 2 shows the results of 1-step 2IFC discrimination, comparing perceptual sensitivity between the Finnish natives and the Dutch students of Finnish. The effect of language exposure on 1-step discrimination was addressed in an ANOVA with between-subjects factors Amount of Exposure (4) and Stimulus Pair (6). Only a main effect of Stimulus Pair was present [$F(5,60) = 6.6, p < .001$].

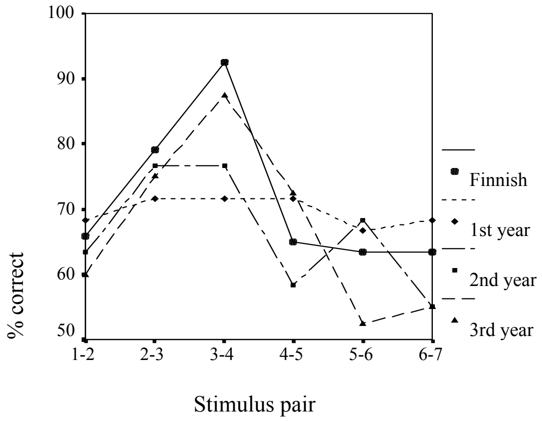

Figure 2: 1-step 2IFC discrimination results of the Finnish natives, and students of Finnish (1st, 2nd, and 3rd year).

Even though overall differences between Finnish and Dutch listeners were absent, Figure 2 shows that the native listeners' discrimination peak is approximated by that of the third-year students. To examine this, the raw data set was examined for each stimulus pair separately in ANOVAs with Amount of Exposure (4) as between-subjects factor. These analyses showed that only responses to Stimulus Pair 3-4, i.e. where the Finnish and third-year students showed increased perceptual sensitivity, differed as a function of amount of exposure [$F(3,276) = 5.6, p = .001$]. Furthermore, post-hoc analyses confirmed that the Finnish listeners' mean percentage of correct responses to Stimulus Pair 3-4 was significantly higher than those of the first- and second-year students, but not higher than those of the third-year students. Apparently, perceptual sensitivity to a nonnative quantity contrast as measured by 2IFC discrimination can become similar to native listeners' sensitivity with ample language exposure.

To study the effect of language exposure on 2-step 2IFC discrimination an ANOVA with between-subjects factors Amount of Exposure (4) and Stimulus Pair (5) was run. A main effect of Stimulus Pair was found [$F(4,50) = 9.2, p < .001$]. The four listener groups all showed high performance levels on this test – group means varied from 85.8% to 89.3% – but could not be distinguished on the basis of their results.

3.3. Conclusion

The present experiment aimed at testing whether Dutch listeners can obtain native-like behavior in perceiving a nonnative quantity contrast. Despite the small number of participants, the results were promising. The predictions that advanced learners will increasingly resemble native Finnish listeners, and that they will do so through acquired distinctiveness, were borne out by the results.

4. General conclusion

The perceptual development of Dutch listeners learning the Finnish quantity contrast /t-tː/ was studied. The question was whether nonnative perception develops through acquired distinctiveness or through acquired similarity. Experiment I showed that Dutch learners of Finnish needed only a few days of training to improve their classification of words containing either /t/ or /tː/. However, they needed much more experience to achieve native-like sensitivity as reflected by a peak near the perceived phoneme boundary along the /t-tː/ continuum. The results of experiment II showed that, with ample language experience, Dutch listeners attained native-like perception of the Finnish quantity contrast. In addition, as the increase in sensitivity near the perceived phoneme boundary demonstrated, perception developed through acquired distinctiveness.

5. References

[1] Laver, J., *Principles of phonetics*, Cambridge University Press, 1994.

[2] Rietveld, T., Kerkhoff, J., and Gussenhoven, C., "Word prosodic structure and vowel duration in Dutch", *Journal of Phonetics*, 32, 349-371, 2004.

[3] Rietveld, A.C.M. and Van Heuven, V.J., *Algemene fonetiek*. Bussum: Coutinho, 1997.

[4] Nooteboom, S.G., *Production and Perception of Vowel Duration*. Doctoral Dissertation, Utrecht University, 1972.

[5] Heuven, V.J. van, "Linguistic versus phonetic explanation of consonant lengthening after short vowel: a contrastive study of Dutch and English", *Proceedings of ICSLP*, 1275-1277, 1992.

[6] Jongman, A. and Sereno, J.A., "On vowel quality and post-vocalic consonant duration in Dutch", *Proceedings of XIIth ICPhS*, 294-297, 1991.

[7] Nooteboom, S.G. and Cohen, A., *Spreken en Verstaan: een Nieuwe Inleiding tot de Experimentele Fonetiek*. Assen: Van Gorcum, 1988.

[8] Liberman, A. M., Harris, K.S., Kinney, J.A. and Lane, H., "The discrimination of relative onset-time of the components of certain speech and nonspeech patterns", *Journal of Experimental Psychology* ***61***: 379-388, 1961.

[9] Heeren, W., "Perceptual development of a new phoneme contrast by adult and 12-year-old listeners", *Proceedings of From Sound to Sense, MIT,* C31-C36, 2004.

[10] Lively, S.E., Logan, J.S. and Pisoni, D.B., "Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories", *Journal of the Acoustical Society of America 94*, 1242-1255, 1993.

[11] Jamieson, D.G. and Morosan, D.E., "Training non-native speech contrasts in adults: Acquisition of the English /ð/-/θ/ contrast by francophones", *Perception & Psychophysics 40*, 205-215, 1986.