Dutch listeners' perception of Korean fortis, lenis, and aspirated stops: First exposure

Mirjam Broersma

Max Planck Institute for Psycholinguistics, Nijmegen, & Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands

mirjam@mirjambroersma.nl

ABSTRACT

This study addresses Dutch listeners' perception of Korean fortis, lenis, and aspirated stop triplets. The Korean stop triplets were expected to be extremely difficult for Dutch listeners, as all Korean stops are voiceless (in initial position), while Dutch only distinguishes voiced and voiceless stops. Dutch listeners were not familiar with Korean and received no training; they only heard six examples of each target sound before performing a phonetic categorization task.

The Korean stops were lenis, fortis, and aspirated bilabial (/p/-/p*/-/p^h/), denti-alveolar (/t/-/t*/-/t^h/), and velar (/k/-/k*/-/k^h/) stops. Stimuli were blocked by type and place of articulation; thus one block concerned either a fortis-lenis, a lenis-aspirated, or a fortis-aspirated contrast, and either bilabial, denti-alveolar, or velar stops.

Both Korean and Dutch listeners performed best on fortis-aspirated, intermediate on lenis-fortis, and poorest on aspirated-lenis contrasts. Not surprisingly, Dutch listeners performed less accurately than Korean listeners. Remarkably, however, the Dutch listeners performed significantly above chance level (if only just above it numerically) for each contrast type at each place of articulation. Thus, even at first exposure and without any training, Dutch listeners managed to identify the extremely difficult L2 sounds relatively successfully.

Keywords: Speech perception; Consonants; Identification; Non-native language; Korean.

1. INTRODUCTION

The identification of phonemes in a second language is one of the greatest challenges for non-native listeners (see e.g. the collected papers in Strange, 1995, and Bohn & Munro, 2007). The largest perceptual difficulties have been proposed to arise in cases in which the second language (L2) has two phonemes where the native language (L1) has only one, as described by the Perceptual Assimilation Model (Best, 1994; Best & Tyler, 2007). In the present study, another extremely problematic case is investigated: a case in which the L2 has *three* phonemes where the L1 has only one phoneme in the same perceptual space. This study investigates Dutch listeners' perception of Korean stop triplets. Dutch listeners were not familiar with Korean, and received no training; they only heard six examples of each target sound before performing phonetic categorization of these truly New Sounds.

Korean has a three-way stop contrast, and all of those stops are (at least in initial position) voiceless. The three stop categories are lenis, fortis, and aspirated, all occurring at bilabial (/p/-/p*/-/p^h/), denti-alveolar (/t/-/t*/-/t^h/) and velar (/k/-/k*/-/k^h/) places of articulation. Dutch, on the other hand, distinguishes prevoiced and voiceless unaspirated stops (/b/-/p/, /d/-/t/, /k/; /g/ only occurs in loanwords). Dutch listeners are likely to perceive all Korean lenis, fortis, and aspirated stops as most similar to Dutch voiceless stops. The Korean three-way distinction can therefore be expected to be very difficult to distinguish for Dutch listeners.

Some of the perceptual cues that are important for the Korean three-way distinction, i.e., the voice quality of the following vowel (with creaky voice after fortis stops, breathy voice after lenis, and breathy or modal voice after aspirated stops; Cho, Jun, & Ladefoged, 2002; Kang & Guion, 2006), do not play a role in Dutch stop voicing contrasts. Others, like VOT and F0 at onset of the following vowel (Cho et al., 2002; Kang &

Guion, 2006; Lisker & Abramson 1964) are also important cues for the stop voicing contrast in Dutch (Van Alphen & Smits, 2004). Those cues, however, have different critical values in the two languages.

This study addresses Dutch listeners' perception of Korean lenis, fortis, and aspirated stops in nonword-initial position at first exposure, and compares their performance to that of native Korean listeners. The study assesses whether Dutch listeners can identify Korean stops at a level above chance, even at first exposure and without receiving any training, and whether Dutch and Korean listeners differ in which stop categories they find most confusable.

2. METHOD

2.1. Participants

Participants were 36 native listeners of Korean and 36 native listeners of Dutch. Korean participants were students at Hanyang University (Seoul, Korea), and Dutch participants at the Radboud University Nijmegen (The Netherlands), who participated for course credits or a small financial compensation. None reported any hearing loss. None of the Dutch participants had any knowledge of or experience with Korean.

2.2. Materials

Target sounds were the nine stops /p/, /p*/, /p^h/, /t/, /t*/, /t^h/, /k/, /k*/, /k^h/. Each target sound occurred in initial position, in three phonetic contexts, followed by the vowel /i/, /u/, or /ɛ/. (Note that /ɛ/ rather than /a/ was used because the vowel /a/ was the target sound in a vowel contrast tested in the same perception experiments as the stops, the results of which are not discussed in this paper.) There were thus 27 items (9 target stops * 3 vowel contexts). Each item was recorded 20 times, yielding a total of 540 stimuli.

The materials were recorded by a 23 year old female native speaker of Korean, who had been born and raised in Seoul. She read the items, presented in Korean orthography, one by one, separated by a pause, in a clear citation style.

The recording was made in a sound proof booth with a Sennheiser microphone and stored directly onto a computer at a sample rate of 41.5 kHz. Stimuli were excised from the recording using the speech editor Praat.

Additionally, Korean vowel, fricative, and affricate materials were used, the results of which are not reported here. Those materials consisted of recordings by the same speaker of 5 Korean vowels, 3 fricatives, and 2 affricates, with 60 stimuli per phoneme.

2.3. Design

Stimuli were blocked by type and place of articulation; thus one block concerned either a fortis-lenis, or a lenis-aspirated, or a fortis-aspirated contrast, and either bilabial, or denti-alveolar, or velar stops. Participants were equally distributed over three stimulus lists. Each list contained one fortis-lenis, one lenis-aspirated, and one fortis-aspirated stop contrast, each for one place of articulation. Thus, no participant heard one type of stop contrast in more than on block, and no participant heard one place of articulation in more than block.

Additionally, participants categorized three other contrasts that are not described here, concerning Korean vowels, fricatives, and affricates.

The order of blocks was counterbalanced, and stimuli within a block were presented in a semi-random order which was different for each participant, with each phoneme occurring maximally four times in succession.

2.4. Procedure

Participants were tested one at a time in a quiet room, seated in front of a computer. They received written instructions in their native language that they would hear a series of items containing one of two sounds, numbered 1 and 2. They were instructed to decide on each trial which of the two sounds they had heard, and to indicate their response by pressing the corresponding response button, labeled 1 or 2, as fast and as accurately as possible.

Before each of the six blocks they received further instructions about the two response alternatives and the corresponding response buttons in that block. First, it was indicated on the computer screen if the block concerned vowels or consonants. Next, 6 unique examples were played of each target sound, accompanied by the number 1 or 2 on the screen. Example stimuli were similar to the experimental stimuli, and were grouped by phonetic context (i.e., examples for the /p/ - /p*/ contrast were pa, pa, p*a, p*a, pi, pi, p*i, p*i, pe, pe, p*e, p*e, in that order).

The experiment started with a short practice part in which participants categorized the Korean /i/-/u/ contrast (which was easy to distinguish for Dutch listeners), to familiarize them with the task.

The experiment was controlled with NESU (Nijmegen Experiment Set-Up) software. Stimuli were played binaurally, one at a time, over Sennheiser closed headphones at a comfortable level. Participants responded by pressing one of two buttons, labeled "1" and "2", on a box in front of them. There was no time limit for the responses. At 600 ms after button press, the next stimulus was played.

3. RESULTS AND DISCUSSION

3.1. Korean listeners

Table 1. Korean listeners' results; % correct, mean *d'*, mean $log\ \beta$, one-tailed One-Sample T Test for *d'* > 0, two-tailed One-Sample T Test for $log\ \beta \neq 0$. (Higher values of *d'* indicate higher sensitivity. Negative values of $log\ \beta$ indicate a bias towards the first, and positive values towards the second phoneme in the first column.)

Contrast	% Correct	*d'*	$log\ \beta$	T Test *d'*	T Test $log\ \beta$
Bilabials					
Lenis – Fortis	90.3	3.96	-1.86	t (11) = 5.8 $p < .001$	t (11) = -1.8 $p = .09$
Fortis – Aspirated	94.8	4.06	-0.74	t (11) = 9.8 $p < .001$	t (11) < |1| $p > .1$
Aspirated – Lenis	79.4	1.79	-0.44	t (11) = 10.2 $p < .001$	t (11) = -2.9 $p < .05$
Denti-alveolars					
Lenis – Fortis	89.0	3.16	-2.56	t (11) = 9.7 $p < .001$	t (11) = -2.4 $p < .05$
Fortis – Aspirated	94.0	3.89	0.74	t (11) = 8.7 $p < .001$	t (11) < |1| $p > .1$
Aspirated – Lenis	78.0	1.65	0.19	t (11) = 9.2 $p < .001$	t (11) = 1.5 $p > .1$
Velars					
Lenis – Fortis	85.7	2.68	-1.55	t (11) = 11.5 $p < .001$	t (11) = -1.8 $p > .1$
Fortis – Aspirated	91.3	3.85	-0.64	t (11) = 7.3 $p < .001$	t (11) < |1| $p > .1$
Aspirated – Lenis	85.2	2.83	1.66	t (11) = 5.8 $p < .001$	t (11) = 1.5 $p > .1$

Korean listeners' categorization results are summarized in Table 1. Mean percentages correct per contrast range from 78.0 to 94.8 %.

For each participant and each contrast separately, *d'* was calculated to assess listeners' sensitivity (with a correction for near-perfect sensitivity, MacMillan & Creelman, 1991). A *d'* of 0 indicates that listeners do not treat two phonemes as different; a *d'* of 1 corresponds to 69 % correct, and the effective upper limit of *d'* is 4.65. One-tailed One-Sample T Tests for each contrast showed that *d'* was always significantly larger than 0 (Table 1). Listeners were thus sensitive to each contrast.

As Table 1 shows, percentage correct (and, with one exception, *d'*) was largest for fortis-aspirated, intermediate for lenis-fortis, and smallest for aspirated-lenis contrasts. An ANOVA with *d'* as the dependent variable and Contrast Type (lenis-fortis, fortis-aspirated, aspirated-lenis), Place of Articulation, and Context (following vowel) as independent variables showed a main effect of Contrast Type (F (2, 99) = 21.4, p < .001). To further investigate this effect, similar ANOVAs were done comparing the Contrast Types pairwise. They confirmed that *d'* was significantly larger for fortis-aspirated contrasts than for both lenis-fortis (F (1, 66) = 9.6, p < .01) and for aspirated-lenis contrasts (F (1, 66) = 34.3, p < .001), and that *d'* was larger for lenis-fortis than for aspirated-lenis contrasts (F (1, 66) = 20.5, p < .001). This pattern was not modified by place of articulation, as there were no interactions between Contrast Type and Place of Articulation.

Further, $log\ \beta$ was calculated to assess possible response biases (McNicol, 1972). A $log\ \beta$ of 0 indicates that there is no bias, a negative $log\ \beta$ that there is a bias towards the first phoneme, and a positive $log\ \beta$ that there is a bias towards the second phoneme mentioned in the first column of the table. Two-tailed One-Sample T Tests for each contrast showed that $log\ \beta$ was significantly different from 0 in two cases; for the aspirated-lenis bilabials, there was a bias towards aspirated ('p^h') responses, and for the lenis-fortis denti-alveolars, there was a bias towards lenis ('t') responses.

3.2. Dutch listeners

Dutch listeners' categorization results are summarized in Table 2. Mean percentages correct per contrast range from 55.3 to 78.9 %.

As for the Korean listeners, for each participant and each contrast separately, *d'* was calculated as a measure of sensitivity, and $log\ \beta$ as a measure of bias. One-tailed One-Sample T Tests for each contrast showed that *d'* was always significantly larger than 0 (Table 2). Despite the sometimes very low percentages correct, listeners thus showed sensitivity to each contrast.

Table 2 shows that, similar to the Korean listeners' results, percentage correct and *d'* were largest for fortis-aspirated, intermediate for lenis-fortis, and smallest for aspirated-lenis contrasts. Like for the Korean listeners, an ANOVA with *d'* as the dependent variable and Contrast Type, Place of Articulation, and Context as independent variables showed a main effect of Contrast Type (F (2, 98) = 19.1, p < .001). To further investigate this effect, similar ANOVAs were done comparing the Contrast Types pairwise. The analyses confirmed again that *d'* was significantly larger for fortis-aspirated contrasts than for both lenis-fortis (F (1, 66) = 4.3, p < .05) and for aspirated-lenis contrasts (F (1, 65) = 38.5, p < .001), and larger for lenis-fortis contrasts than for aspirated-lenis contrasts (F (1, 65) = 16.7, p < .001). Like for the Korean listeners' results, this pattern was not modified by place of articulation, as there were no interactions between Contrast Type and Place of Articulation.

Comparing the Dutch and Korean listeners' percentages correct and *d'* values (Tables 1 and 2), it is clear that the Dutch listeners had a lower accuracy than the Korean listeners for all contrasts. Indeed, in an ANOVA with *d'* as the dependent variable and Language Group (Dutch and Korean), Contrast Type, Place of Articulation, and Context as independent variables, there was a main effect of Language Group (F (1, 197) = 149.5, p < .001). Because there were also significant interactions among Language Group, Context, and Contrast Type (F (4, 394) = 2.7, p < .05), and among Language Group, Context, and Place of Articulation (F (4, 394) = 3.1, p < .05), the effect of Language Group was also calculated for each contrast separately. For each contrast, *d'* was significantly larger for the Korean listeners than for the Dutch listeners (Table 2).

Finally, *log β* was calculated as a measure of bias for each participant and each contrast separately. Two-tailed One-Sample T Tests for each contrast showed that *log β* was significantly different from 0 in one case; for the lenis-fortis bilabials, there was a bias towards lenis ('p') responses. This is different from the Korean listeners' results, and indeed, in an ANOVA with *log β* as the dependent variable, there was a significant interaction among Language Group, Contrast Type, and Place of Articulation (F (4, 197) = 2.5, $p < .05$).

Table 2. Dutch listeners' results; % correct, mean *d'*, mean *log β*, one-tailed One-Sample T Test for *d'* > 0, F test for *d'*: main effect of Language Group, two-tailed One-Sample T Test for *log β* ≠ 0. (Higher values of *d'* indicate higher sensitivity. Negative values of *log β* indicate a bias towards the first, and positive values towards the second phoneme in the first column.)

Contrast	**% Correct**	***d'***	***log β***	**T Test *d'***	**F test *d'*, main effect of Language Group**	**T Test *log β***
Bilabials						
Lenis – Fortis	67.9	1.12	-0.54	t (11) = 5.6 $p < .001$	F (1, 22) = 19.4 $p < .001$	t (11) = -2.9 $p < .05$
Fortis – Aspirated	78.9	2.24	0.55	t (11) = 4.3 $p < .001$	F (1, 22) = 8.5 $p < .01$	t (11) < |1| $p > .1$
Aspirated – Lenis	55.3	0.28	-0.03	t (11) = 2.2 $p < .05$	F (1, 21) = 45.8 $p < .001$	t (11) < |1| $p > .1$
Denti-alveolars						
Lenis – Fortis	66.1	0.96	0.14	t (11) = 3.9 $p < .01$	F (1, 22) = 35.0 $p < .001$	t (11) < |1| $p > .1$
Fortis – Aspirated	76.0	2.15	0.16	t (11) = 3.6 $p < .01$	F (1, 22) = 9.7 $p < .01$	t (11) < |1| $p > .1$
Aspirated – Lenis	56.1	0.40	-0.25	t (11) = 2.2 $p < .05$	F (1, 22) = 22.4 $p < .001$	t (11) = -1.4 $p > .1$
Velars						
Lenis – Fortis	63.0	0.77	-0.24	t (11) = 4.3 $p < .001$	F (1, 22) = 31.2 $p < .001$	t (11) = -1.8 $p > .1$
Fortis – Aspirated	76.7	1.89	-0.50	t (11) = 4.4 $p < .001$	F (1, 22) = 8.9 $p < .01$	t (11) < |1| $p > .1$
Aspirated – Lenis	55.6	0.29	-0.01	t (11) = 2.6 $p < .05$	F (1, 22) = 25.3 $p < .001$	t (11) < |1| $p > .1$

4. GENERAL DISCUSSION

The results showed that, as expected, Dutch listeners found it very difficult to distinguish the Korean three-way stop contrasts. For each of the contrast types at each place of articulation, the Dutch listeners were significantly less accurate than the Korean listeners. Despite percentages of correct responses that were sometimes as low as 55-56 %, the Dutch listeners nevertheless performed significantly above chance level for all contrast types at all places of articulation; their sensitivity as measured by *d'* was always significantly above 0.

Dutch and Korean listeners showed some differences in response biases. Importantly, however, they did not differ in which combinations of stop types they found most difficult. Both Dutch and Korean listeners performed best on the identification of fortis versus aspirated stops, intermediate on lenis versus fortis stops, and worst on aspirated versus lenis stops, at all places of articulation. This might be because the perceptual cues that have been shown to be most important for Korean listeners' recognition of the three-way stop contrast, i.e., VOT and F0 of the following vowel (Cho et al., 2002; Kang & Guion, 2006), also play a role for Dutch stop voicing perception (Van Alphen & Smits, 2004). Thus, even though Dutch listeners were not familiar with the Korean stops, they might have attempted to use the same perceptual cues as the Korean listeners did to some extent. Saliency of those cues might have resulted in a similar pattern from difficult to less difficult contrasts for Dutch and Korean listeners alike.

Thus, even at first exposure and without any training, Dutch listeners managed to identify the Korean fortis, lenis, and aspirated stops at a level above chance. As Korean stop triplets are arguably an example of the worst possible L2 contrasts for Dutch listeners, with three L2 phonemes in the same perceptual space where the L1 has only one phoneme, it is remarkable that the Dutch listeners managed to identify the sounds successfully, even if that success only meant that they recognized them with an accuracy that was just above chance level.

5. REFERENCES

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (Eds.), *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words* (pp. 167-224). Cambridge, MA: MIT.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege* (pp. 13-34). Amsterdam: John Benjamins.

Bohn, O.-S., & Munro, M. J. (Eds.). (2007). *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*. Amsterdam: John Benjamins.

Cho, T., Jun, S.-A., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics, 30*, 193-228.

Kang, K.-H., & Guion, S. G. (2006). Phonological systems in bilinguals: Age of learning effects on the stop consonant systems of Korean-English bilinguals. *Journal of the Acoustical Society of America, 119*, 1672-1683.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*, 384-422.

MacMillan, N. A., & Creelman, C. D. (1991). *Detection Theory: A User's Guide*. Cambridge: CUP.

McNicol, D. (1972). *A Primer of Signal Detection Theory*. Sydney: Australasian Publishing Company.

Strange, W. (Ed.). (1995). *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Baltimore: York Press.

Van Alphen, P. M., & Smits, R. (2004). Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: The role of prevoicing. *Journal of Phonetics, 32*, 455-491.