LETTERS TO THE EDITOR

This Letters section is for publishing (a) brief acoustical research or applied acoustical reports, (b) comments on articles or letters previously published in this Journal, and (c) a reply by the article author to criticism by the Letter author in (b). Extensive reports should be submitted as articles, not in a letter series. Letters are peer-reviewed on the same basis as articles, but usually require less review time before acceptance. Letters cannot exceed four printed pages (approximately 3000–4000 words) including figures, tables, references, and a required abstract of about 100 words.

Flexible cue use in nonnative phonetic categorization (L)

Mirjam Broersma[a)]
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands, Radboud University, Nijmegen, 6500 HE The Netherlands

(Received 1 August 2007; revised 8 May 2008; accepted 9 May 2008)

Native and nonnative listeners categorized final /v/ versus /f/ in English nonwords. Fricatives followed phonetically long (originally /v/-preceding) or short (originally /f/-preceding) vowels. Vowel duration was constant for each participant and sometimes mismatched other voicing cues. Previous results showed that English but not Dutch listeners (whose L1 has no final voicing contrast) nevertheless used the misleading vowel duration for /v/-/f/ categorization. New analyses showed that Dutch listeners did use vowel duration initially, but quickly reduced its use, whereas the English listeners used it consistently throughout the experiment. Thus, nonnative listeners adapted to the stimuli more flexibly than native listeners did. © *2008 Acoustical Society of America.*
[DOI: 10.1121/1.2940578]

PACS number(s): 43.71.Hw, 43.71.Es, 43.71.Sy [AJ] Pages: 712–715

I. INTRODUCTION

This paper investigates whether nonnative listeners can be more flexible than native listeners in their use of perceptual cues for making phoneme distinctions. In a previous study (Broersma, 2005), Dutch and English listeners categorized English obstruents at the end of nonwords as voiced or voiceless. Like English, Dutch has an obstruent voicing distinction, but unlike English, it has no voiced obstruents at the end of words in isolation (Booij, 1995). Thus, although Dutch listeners are familiar with the voicing distinction from their native language, they are not familiar with making this distinction in final position, or with the use of preceding vowel duration as a cue for this distinction. Preceding vowel duration is a perceptual cue for word-medial voicing distinctions in Dutch (Van den Berg, 1989), but a less prominent one than for the final voicing distinction in English (e.g., Raphael, 1972). Dutch listeners seem to be able to generalize their knowledge about vowel duration and word-medial obstruent voicing to word-final obstruents, as they categorize the duration of vowels differently depending on the underlying voicing of Dutch word-final obstruents (Jongman *et al.*, 1992).

Previous results (Broersma, 2005) showed that Dutch listeners accurately distinguished English final voicing contrasts, but that they did not use vowel duration in the same way English listeners did. In an experiment where vowel duration was made uninformative and even misleading, English listeners persisted in using this cue for making final /v/-/f/ distinctions, whereas Dutch listeners did not use the cue significantly. Vowel duration was made uninformative to preclude the possibility that the Dutch listeners used vowel duration only as a result of the nature of the stimulus materials. This leaves open the possibility, however, that Dutch listeners did *not* use vowel duration only as a result of the nature of the materials. Thus, the Dutch listeners might have more flexibly adapted to the information in the stimulus materials than the English listeners.

In the Broersma (2005) experiment, an 11-step /v/ to /f/ continuum was combined with a phonetically long and a phonetically short vowel. However, each participant heard only one of the vowels throughout the experiment. Thus, vowel duration was not informative and, moreover, mismatched the voicing information in the fricatives for some of the items, as v-like fricatives were sometimes preceded by a short vowel and f-like fricatives by a long vowel. In order to assess the use of vowel duration as a cue for final fricative voicing, 50-percent crossover points in the conditions with long and short preceding vowels were compared. For the English listeners, there was a significant shift in the 50-percent crossover point, with more "voiced" responses in the Long Vowel condition than in the Short Vowel condition. For the Dutch listeners, there was a small trend in the same direction, but this was not significant. Further, the Dutch listeners' categorization curve was steeper than the English listeners' curve in the Long Vowel condition. (There was no difference in the Short Vowel condition.)

In the same experiment, for a final /z/ to /s/ continuum, no shift in the 50-percent crossover points in the Long Vowel versus the Short Vowel condition was found, for the English

a)Electronic mail: mirjam.broersma@mpi.nl.

 0001-4966/2008/124(2)/712/4/$23.00

TABLE I. The 50-percent crossover point and steepness of the categorization curve at that point (with lower values indicating steeper slopes), for Dutch and English listeners, in the Long Vowel and Short Vowel conditions.

	50-percent crossover point				Steepness			
	Dutch		English		Dutch		English	
Part	Long	Short	Long	Short	Long	Short	Long	Short
1	6.7	5.8	6.9	5.0	0.4	0.4	0.5	0.2
2	6.4	5.9	7.7	5.9	0.3	0.4	0.4	0.3
3	6.3	5.5	6.9	5.8	0.4	0.3	0.5	0.3
4	6.4	5.8	7.2	5.3	0.3	0.3	0.5	0.3
5	6.4	6.0	7.0	5.4	0.3	0.4	0.5	0.3
6	6.2	6.0	6.0	5.0	0.4	0.3	0.4	0.3
1–6	6.4	5.8	7.0	5.4	0.4	0.3	0.5	0.3

or for the Dutch listeners. There are spectral differences between the alveolar and the labiodental stimuli that may explain the lack of such a shift for the /z/ to /s/ continuum. First, the difference in intensity between voiced and voiceless fricatives in the frequency range important for voicing was larger for the alveolar than for the labiodental fricatives (intensity of the first spectral peak below 500 Hz, measured from 10 ms at the center of the fricative, with means of fast Fourier transform using a Gaussian window: /z/-/s/: 29.5 −9.5=20.0 dB; /v/-/f/: 24.7−10.9=13.8 dB). Further, the overall intensity of the alveolar fricatives was higher than that of the labiodental fricatives (/z/: 67.7; /s/: 64.0; /v/: 65.8; /f/: 57.9 dB) (cf. Jongman *et al.*, 2000), and that of the preceding vowels lower (alveolars, long vowel: 76.1, short vowel: 66.5 dB; labiodentals, long vowel: 77.9, short vowel: 69.0 dB). Thus, as the alveolar fricatives contained larger amplitude differences related to voicing, and their frication was (both absolutely and relatively to the vowel) more easily perceptible than that of the labiodental fricatives, listeners may have been less inclined to exploit vowel duration as a cue to voicing for the alveolar than for the labiodental fricatives.

For the /v/ to /f/ continuum, the Dutch listeners might have discovered during the experiment that vowel duration was not a helpful cue to final fricative voicing here and learned to ignore it. This paper attempts to test this explanation by establishing whether the listeners did not use vowel duration from the start of the experiment, or tried to use vowel duration as a voicing cue initially, but stopped doing so at some point in the experiment because of the nature of the stimulus materials. Relevant data could be obtained from the results of the practice part of the study, which were not taken into account in Broersma (2005).

II. METHOD

In a Two-Alternative Forced Choice experiment, listeners categorized sounds as "v" or "f." Participants were 28 native speakers of English, and 28 native speakers of Dutch who were proficient in English as a second language.

The materials consisted of 11 fricatives spliced onto two carriers. The carriers were one token of /ku:/ with a phonetically long vowel (extracted from a recording of /ku:v/), and one with a phonetically short vowel (extracted from a recording of /ku:f/). The fricatives formed a continuum from a natural /v/ to a natural /f/ (extracted from another recording of /ku:v/ and /ku:f/) with nine intermediate steps, created following the procedure of Stevenson (1979) and Repp (1981) by adding up the waveforms of the /v/ and the /f/ in varying, equally spaced proportions.

In the practice part, participants categorized each step of the continuum three times, then there was a break during which they could ask questions, and finally they categorized each step one more time. After that, the main part of the experiment, consisting of 20 presentations of each step, started without any further demarcation. All items, in the practice part as well as the main part of the experiment, were presented in semirandomized order. Items in the Long Vowel condition contained the phonetically long vowel and those in the Short Vowel condition the phonetically short vowel. Participants were randomly assigned to a condition, with equal numbers in both conditions. Each participant thus heard only the long vowel or only the short vowel throughout the entire experiment. For more details about the method, see Broersma (2005).

III. RESULTS

Each subject's categorization curve was fitted with logistic regression to determine the 50-percent crossover point and the steepness of the curve at that point. First, the results of the practice part and the main part of the experiment were analyzed together (Table I, part 1-6), with an analysis of variance (ANOVA) with 50-percent crossover point as the dependent variable and vowel duration and native language as independent variables. There was a significant interaction between vowel duration and native language ($F(1,41)=5.67, p<0.05$), showing that the effect of vowel duration was larger for the English listeners than for the Dutch listeners. However, the main effect of vowel duration was significant not only for the English listeners ($F(1,18)=32.51, p<0.001$), but also for the Dutch listeners ($F(1,23)=4.74, p<0.05$). Thus, in contrast to Broersma (2005), when the practice part was included, Dutch listeners showed significant evidence of the use of vowel duration for final fricative voicing decisions.

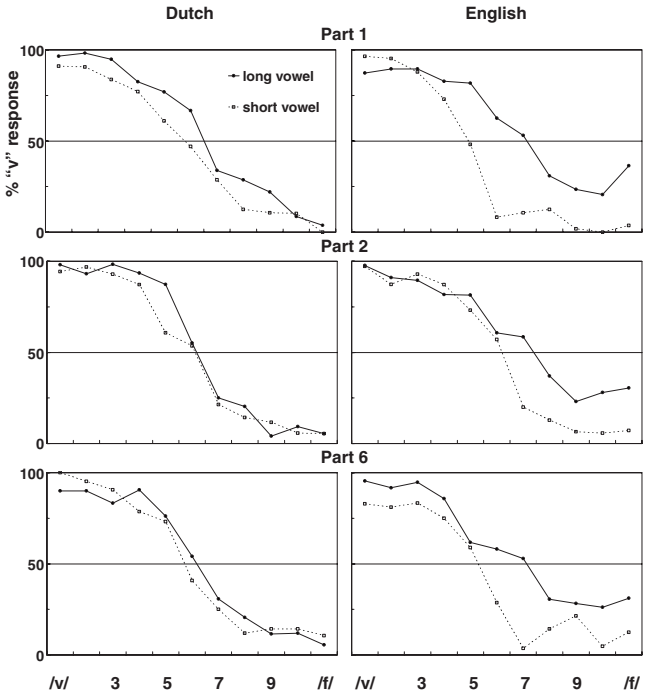

FIG. 1. Mean percentage of "v" responses as a function of the place on an 11-step /v/ to /f/ continuum and preceding vowel duration, per experimental part (1, 2, and 6) and participants' native language.

Next, the practice part was analyzed separately (Fig. 1 and Table I, part 1). There was a main effect of vowel duration ($F(1,47)=15.15, p<0.001$), no main effect of native language ($F(1,47)<1$), and no interaction between native language and vowel duration ($F(1,47)=1.84, p>0.1$). Thus, in the practice part, Dutch and English listeners' responses depended similarly on the duration of the preceding vowel.

Finally, the change in the use of vowel duration over time, in particular from the practice part to the main part of the experiment, was assessed. The main part of the experiment was split into parts of the same length as the practice part (44 trials). The practice part was called Part 1, and the main part was divided into Parts 2–6, with Part 2 containing the first 44 trials of the main part, Part 3 the next 44 trials, and so on. Part 6 contained only 33 trials. There was a three-way interaction among Part, vowel duration, and language ($F(5,205)=2.40, p<0.05$). Crucially, comparing the results of the practice (Part 1) to those of Part 2 (Fig. 1 and Table I, part 2), for the Dutch listeners, there was an interaction between Part and vowel duration ($F(1,25)=4.46, p<0.05$). For these listeners, the effect of vowel duration was significant in Part 1 ($F(1,25)=5.63, p<0.05$), but not in Part 2 ($F(1,26)=1.2, p>0.1$). For the English listeners, there was no interaction between Part and vowel duration for Part 1 versus 2 ($F(1,21)=1.97, p>0.1$). Comparing all consecutive Parts in a similar, pairwise manner (i.e., Part 2 with 3, 3 with 4, 4 with 5, and 5 with 6), there were no other interactions between Part and vowel duration for the Dutch or for the English listeners. Thus, for the Dutch listeners, the use of vowel duration decreased from Part 1 to Part 2, and then stayed the same until the end of the experiment, while for the English listeners the use of vowel duration did not change throughout the experiment.

To assess how categorical the listeners' perceptual performance was, ANOVAs on the steepness of the categorization curves were performed. Recall that in the main part of the experiment, the Dutch listeners' categorization curve was steeper than the English listeners' curve in the Long Vowel condition, and there was no difference in the Short Vowel condition. In the practice part, there was also an interaction between native language and vowel duration ($F(1,47)=5.88, p<0.05$). Here, however, the English listeners' categorization curve was steeper than the Dutch listeners' curve in the Short Vowel condition ($F(1,24)=6.54, p<0.05$), and there was no difference in the Long Vowel condition ($F(1,23)<1$). Comparing Part 1 and Part 2, there was an interaction between Part and native language ($F(1,46)=5.40, p<0.05$), and there were no such interactions for any other set of consecutive Parts. Thus, whereas the Dutch listeners' perception was less categorical than the English listeners' perception in the practice part, it was more categorical than the English listeners' responses in the main part of the experiment.

IV. DISCUSSION

The analysis of the practice part shows that the Dutch listeners initially used vowel duration as a cue for final /v/-/f/ categorization to the same extent as English listeners did. However, their use of it rapidly diminished over time. The English listeners, on the other hand, did not change their use of vowel duration as a voicing cue throughout the experiment. The English listeners, who initially categorized (some of) the fricatives more categorically, later categorized (some of) them less categorically than the Dutch listeners: In the practice part they perceived the fricatives in the Short Vowel condition more categorically than the Dutch listeners did, whereas in the rest of the experiment they perceived the fricatives in the Long Vowel condition less categorically than the Dutch listeners did. This change occurred at the same time that the Dutch listeners changed their use of vowel duration (between the practice part and the first 44 trials of the main part of the experiment), and may be related to the Dutch listeners' decreased use of this misleading cue and the English listeners' persistent use of it. This confirms the conclusion from Broersma (2005) that the nonnative listeners did not use vowel duration as a voicing cue in the same way the English listeners did. It seems, however, that they were no less capable of using it, but that they adapted to the nature of the stimulus materials better than the native listeners did. Generally, nonnative listeners might be less certain about which perceptual cues to use than native listeners are. Further, the Dutch listeners' limited experience with the use of vowel duration as a voicing cue might have made it easier for them to ignore this cue than for the native listeners, with their extensive experience with it.

While insecurity about the perceptual relevance of phonetic information and less practice with a perceptual cue might sometimes hinder speech perception, in this study, the nonnative listeners' greater flexibility proved to be an advantage over the native listeners' firmer and presumably more secure cue weighting strategies. For nonnative listeners, who

are trying to work out how phonetic information indicates phonological distinctions in a particular second language, which may be very different from their native language, a large degree of flexibility indeed seems useful.

ACKNOWLEDGMENTS

Many thanks to Sarah Schimke for suggesting these analyses, to Anita Wagner for help with spectral measurements, and to two anonymous reviewers for helpful comments.

Booij, G. (**1995**). *The Phonology of Dutch* (Oxford University Press, Oxford).

Broersma, M. (**2005**). "Perception of familiar contrasts in unfamiliar positions," J. Acoust. Soc. Am. **117**, 3890–3901.

Jongman, A., Sereno, J. A., Raaijmakers, M., and Lahiri, A. (**1992**). "The phonological representation of [voice] in speech perception," Lang Speech **35**, 137–152.

Jongman, A., Wayland, R., and Wong, S. (**2000**). "Acoustic characteristics of English fricatives," J. Acoust. Soc. Am. **108**, 1252–1263.

Raphael, L. J. (**1972**). "Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English," J. Acoust. Soc. Am. **51**, 1296–1303.

Repp, B. H. (**1981**). "Perceptual equivalence of two kinds of ambiguous speech stimuli," Bull. Psychon. Soc. **18**, 12–14.

Stevenson, D. C. (**1979**). *Categorical Perception and Selective Adaptation Phenomena in Speech*. Doctoral dissertation, University of Alberta, Edmonton, Canada.

Van den Berg, R. J. H. (**1989**). "Perception of voicing in Dutch two-obstruent sequences: Covariation of voicing cues," Speech Commun. **8**, 17–25.