

# Perception of final fricative voicing: Native and nonnative listeners' use of vowel duration

Mirjam Broersma<sup>a)</sup>

*Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, 6500 HE Nijmegen, The Netherlands*

(Received 11 September 2008; revised 13 October 2009; accepted 22 December 2009)

Does experience with a perceptual cue for a phoneme contrast in the native language affect its use in a second language for a similar contrast in a different phonetic context? Two experiments investigated Dutch and English listeners' use of preceding vowel duration as a perceptual cue for nonword-final fricative voicing in English. Dutch listeners have native language experience with the use of vowel duration for vowel length and intervocalic obstruent voicing contrasts, but not for final voicing contrasts, as Dutch does not have voiced obstruents word-finally. Previous research [Broersma, M. (2005). *J. Acoust. Soc. Am.* **117**, 3890–3901; (2008) **124**, 712–715] showed that Dutch listeners used vowel duration less for final /v-f/ categorization than English listeners did when vowel duration varied only between subjects, discouraging its use as a perceptual cue. The present study assessed the use of vowel duration for final /v-f/ and /z-s/ contrasts when it varied within subjects. A goodness rating and a phonetic categorization experiment showed that Dutch listeners used vowel duration, but less than English listeners did. Thus, experience with a perceptual cue for a different contrast and for a similar contrast in a different position in the native language did not lead to native-like use of this cue in the second language.

© 2010 Acoustical Society of America. [DOI: 10.1121/1.3292996]

PACS number(s): 43.71.Hw, 43.71.Es, 43.71.Sy [ADP]

Pages: 1636–1644

## I. INTRODUCTION

A lot is known about the perception of the sounds of a second language. A large body of research (see, e.g., the papers in [Strange, 1995](#) and [Bohn and Munro, 2007](#)) has shown that mismatches between the phoneme repertoires of the first (L1) and second (L2) languages cause perceptual difficulties and that, therefore, distinguishing some L2 speech sounds is extremely difficult ([Best and Strange, 1992](#)), whereas distinguishing others is easy ([Best et al., 1988](#)). Such research findings have inspired models of L2 or foreign language speech sound perception like the speech learning model ([Flege, 1995](#)) and the perceptual assimilation model (PAM) ([Best et al., 1988](#); [Best and Tyler, 2007](#)).

Not much is known, on the other hand, about the use of perceptual cues in L2 sound perception. It is not clear whether experience with the use of a perceptual cue in the L1 affects its use in the L2. In other words, can listeners “transfer” the use of an L1 perceptual cue to perception in the L2?

Experience with perceptual cues from the L1 might affect perception of similar L2 contrasts (e.g., French listeners' experience with preceding vowel duration as a cue for the final /z-s/ contrast may have affected their use of this cue for final /z-s/ decisions in English too; [Flege and Hillenbrand, 1986](#)) but it might also aid perception of completely different contrasts. [Crowther and Mann \(1992\)](#) found that native listeners of Japanese, which distinguishes between long and short vowels, were also sensitive to vowel duration when they categorized English final /d/ and /t/. Their experience

with vowel duration as a perceptual cue for Japanese vowel contrasts might thus have aided their perception of English voiced and voiceless final stops.

Experience with perceptual cues from the L1 may, however, not always affect perception of L2 contrasts, even when the perceptual cue is used for similar L1 and L2 contrasts but in a different phonetic position. [Broersma \(2005, 2008\)](#) found that native listeners of Dutch, which has /v-f/ and /z-s/ contrasts similar to the English contrasts but does not distinguish between voiced and voiceless obstruents in word-final position, used the duration of the preceding vowel less consistently than English listeners did. Dutch listeners have native language experience with the use of vowel duration as a perceptual cue for vowel length contrasts and for intervocalic consonant voicing; yet, they did not seem to apply this experience when categorizing English final fricatives.

This paper further investigates the transfer of the use of perceptual cues from the L1 to the L2. It investigates to which extent Dutch listeners, who are familiar with the use of vowel duration as a perceptual cue, but not for final voicing contrasts (as their L1 does not have those), use vowel duration for this purpose when listening to English. To that end, the study of [Broersma \(2005, 2008\)](#) was extended with a different experimental design and a different paradigm.

In English, vowels are generally longer before a voiced final fricative than before a voiceless final fricative ([Jones, 1950](#)), and the duration of the preceding vowel affects listeners' perception of final fricatives as voiced or voiceless ([Denes, 1955](#)). Duration of the preceding vowel is a robust and very important perceptual cue for final fricative voicing for English listeners ([Derr and Massaro, 1980](#); [Hogan and Rozsybal, 1980](#); [Raphael, 1972](#)).

<sup>a)</sup>Electronic mail: mirjam@mirjambroersma.nl

Like English, Dutch has voiced and voiceless obstruents, namely, /b-p/, /d-t/, /v-f/, /z-s/, and /k/. All of the voiceless and none of the voiced obstruents can occur at the end of words in isolation (Booij, 1995). Dutch thus has a /z-s/ and a /v-f/ contrast, but does not have those contrasts in word-final position. Dutch listeners nevertheless do have some native language experience with the use of vowel duration for fricative voicing. For intervocalic two-obstruent sequences in Dutch (with the first, the second, or both being a fricative), duration of the preceding vowel plays a small but significant role for voicing perception (Van den Berg, 1987, 1988, 1989). For single intervocalic fricative voicing, no effect of vowel duration has been demonstrated on perception (e.g., Slis and Cohen, 1969b; Slis and Van Heugten, 1989), but vowels preceding a voiced intervocalic fricative are on average 40 ms longer than those preceding a voiceless intervocalic fricative in Dutch (Slis and Cohen, 1969a). Note that this difference for intervocalic fricatives in Dutch is much smaller than that before final fricatives in English; e.g., Peterson and Lehiste (1960) showed an average difference of 148 ms before voiced and voiceless final fricatives. Finally, results by Jongman *et al.* (1992) suggest that Dutch listeners can generalize their knowledge about the relation between vowel duration and intervocalic obstruent voicing to the case of word-final voicing, as the underlying voicing of voiceless word-final obstruents affects the way Dutch listeners categorize the duration of preceding vowels. Further, Dutch also has phonemically long and short vowels that are mainly distinguished by phonetic vowel duration (Booij, 1995). Thus, Dutch listeners have native language experience with the use of vowel duration as a perceptual cue both for obstruent voicing and vowel length contrasts.

The Dutch listeners studied by Broersma (2005, 2008) had a high level of proficiency in English as a second language and ample exposure to English. They had started receiving English instruction in primary school and were regularly exposed to English, i.e., through the media and in the educational system, as is common in The Netherlands. They therefore seem well equipped for a native-like use of preceding vowel duration as a perceptual cue for the English final /v-f/ and /z-s/ contrasts. Yet, the results of Broersma (2005, 2008) showed robust differences between the Dutch and the English listeners' use of vowel duration as a final fricative voicing cue in English.

In the experiment of Broersma (2005, 2008), vowel duration was made uninformative, and whereas the Dutch listeners initially used vowel duration to the same extent as the English listeners did, they soon stopped using it, whereas the English listeners kept using it persistently throughout the experiment. Thus, the Dutch but not the English listeners adapted their use of vowel duration to the nature of the experimental materials. Vowel duration was made uninformative to preclude the possibility that the Dutch listeners would use vowel duration only as a result of the stimulus materials, as in experiments with a limited number of variables, listeners may rely more on those variables than they would do with normal speech (Bohn, 1995; Bradlow, 1996). Thus, each participant heard either a phonetically long or a phonetically short vowel throughout the experiment, combined

with a fricative from an 11-step voiced to voiceless continuum. Because for each participant vowel duration was kept constant, it was not an informative cue for final fricative voicing. Moreover, vowel duration mismatched the voicing information in the fricatives for some of the items (i.e., for /v/-like fricatives preceded by a short vowel and for /f/-like fricatives preceded by a long vowel).

Broersma (2005) showed that Dutch listeners used vowel duration less than English listeners did as a perceptual cue for the final /v-f/ contrast. Broersma (2008) further showed that the difference between the Dutch and English listeners' results was not immediately present, but arose during the first part of the experiment. In the 44 trials of the practice part of the experiment, the Dutch and the English listeners used vowel duration to a similar extent. After that, vowel duration did not affect Dutch listeners' responses anymore, whereas its effect remained unchanged for the English listeners' responses. For the English listeners, vowel duration strongly affected the responses, such that listeners in the long vowel condition gave more voiced responses than listeners in the short vowel condition did. Long vowels affected their categorization responses especially strongly; long vowels combined with fricatives at the voiceless end point of the continuum received 31% "V" responses. It was concluded that the Dutch listeners adapted to the nature of the stimulus materials more flexibly than the English listeners, who had extensive native language experience with vowel duration as a cue for final fricative voicing. The difference between the Dutch and the English listeners was thus at least partly related to the fact that vowel duration was kept constant for each participant due to the between-subjects design of the experiment.

Another indication that the experimental design induced a limited use of vowel duration as a perceptual cue for voicing are the categorization results for the final /z-s/ contrast (Broersma, 2005). For this contrast, neither the English nor the Dutch listeners' responses showed an effect of vowel duration, while previous research has shown that vowel duration is an important cue for final voicing distinctions for this contrast at least for English listeners (e.g., Derr and Mas-saro, 1980; Hogan and Rozsypal, 1980; Raphael, 1972). The question thus remains to which extent the results were due to the experimental design, and if there are also differences between Dutch and English listeners' use of vowel duration for the final /v-f/ contrast, and possibly for the final /z-s/ contrast, if vowel duration is not kept constant for each participant.

The present paper investigates Dutch and English listeners' use of vowel duration for final fricative voicing in English, with an experimental design that does not discourage the use of vowel duration as a perceptual cue. The same stimuli are used as in Broersma (2005, 2008), but now, instead of blocking vowel duration per participant, all participants are presented with both vowel durations (and with the 11 steps of the fricative voicing continuum, as in the earlier experiment). In Experiment 1, Dutch and English listeners' goodness ratings of the stimuli are collected. In Experiment 2, Dutch and English listeners' phonetic categorization of the

same stimuli is assessed. Experiment 2 is thus a replication of Broersma (2005, 2008), except that now, crucially, vowel duration is varied within subjects.

Both experiments address three questions. First, is there a difference between the Dutch and the English listeners' use of vowel duration for the /v-f/ contrast when it is an informative perceptual cue within the experiment? Second, is vowel duration used as a perceptual cue for the /z-s/ contrast as well, unlike in Broersma (2005, 2008) and, third, if so, do Dutch and English listeners use vowel duration differentially for the /z-s/ contrast?

## II. EXPERIMENT 1

### A. Method

#### 1. Participants

Participants were 16 native speakers of Dutch and 16 native speakers of British English. None of them had participated in the experiments reported in Broersma (2005), but they were part of the same populations tested there. The Dutch participants were proficient in English as a second language. They had received on average 7 years of English instruction in primary and secondary schools and were regularly exposed to English through the media and at the university they attended. The English participants did not know any Dutch. The Dutch participants were recruited from the Max Planck Institute participant pool, and the English participants from the participant pool of the Laboratory of Experimental Psychology of the University of Sussex. None reported any speech or hearing disorders. All were volunteers and received a small fee for participation.

#### 2. Materials

The same materials were used as in Broersma (2005). A male native speaker of British English recorded two tokens of the nonwords /ku:v/, /ku:f/, /fu:z/, and /fu:s/, in a clear citation style, in a soundproof booth. The materials were recorded with a Sennheiser microphone onto digital audiotape and downsampled to 16 kHz during transfer to a computer. Further editing was done with XWAVES software.

For each nonword, from one token, the final fricative was extracted. Thus, /v/, /f/, /z/, and /s/ were extracted from the end of each nonword, with the cut being made at the last positive zero crossing before the onset of frication noise. Two continua, one from /v/ to /f/ and one from /z/ to /s/, were created following the procedure of Stevenson (1979) and Repp (1981). The continua consisted of 11 steps, ranging from the natural voiced end point /v/ or /z/ (step 1) to the natural voiceless end point /f/ or /s/ (step 11), with 9 intermediate steps. To this end, first, the final /f/ and /s/ were shortened to match the durations of the final /v/ and /z/.<sup>1</sup> They were shortened by 56–127 ms (31%) for the /f/ and by 80–187 ms (30%) for the /s/ by removing a portion from the center of the fricative. Next, for the pairs of phonemes thus obtained, the amplitudes of the waveforms were added in varying proportions in 11 equally spaced steps (Stevenson, 1979; Repp, 1981).

For each nonword, a carrier containing all but the fricative was extracted from the other token. Thus, from the non-

words /ku:v/ and /ku:f/, /ku:/ was extracted, and from the nonwords /fu:z/ and /fu:s/, /fu:/ was extracted, removing the final fricative, truncating the signal at the last positive zero crossing before the start of the frication noise. For each continuum, there were thus two carriers. One carrier was originally pronounced with a voiced final fricative and contained a phonetically long vowel (of 257 ms for the /v-f/ and 233 ms for the /z-s/ contrast) and the other was originally pronounced with a voiceless final fricative and contained a phonetically short vowel (of 98 ms for the /v-f/ and 118 ms for the /z-s/ contrast).

For each continuum, the two end points and the nine intermediate steps were spliced onto the end of the appropriate carriers. Thus, for the /v-f/ continuum, all steps were spliced onto both /ku:/ carriers, and for the /z-s/ continuum, all steps were spliced onto both /fu:/ carriers. Auditory examination ensured that the manipulation had not resulted in audible discontinuities.

#### 3. Design

Stimuli were blocked by contrast but not by vowel duration. There were four blocks, two for the /v-f/ contrast and two for the /z-s/ contrast. For the /v-f/ items, participants decided in one block to which extent each fricative sounded like a good “V,” and in the other block to which extent the same items sounded like a good “F,” and for the /z-s/ items, participants decided in one block to which extent each fricative sounded like a good “Z,” and in the other block like a good “S.” As the carrier for the /z-s/ contrast contained an /f/, the /v-f/ contrast was always tested before the /z-s/ contrast; otherwise the order of the blocks was counterbalanced. Each block contained all combinations of 11 fricatives and two carriers presented in random order.

#### 4. Procedure

Participants were tested one at a time in a quiet room. They received written instructions in their native language that they would hear a series of English nonwords, and they were asked to pay attention to the final sound. Before each block, they were informed about the target sound in that block. Participants were instructed to decide whether the final sound of each stimulus was a very poor, fairly poor, fairly good, or very good instance of the target sound and to indicate their response by pressing one of four buttons, labeled from “1—poor” to “4—good.”

The experiment started with a short practice during which the participants heard nonwords ending with a /b/ or a /p/, some of them unedited and others with the release burst removed (Broersma 2005, Experiment 1) and indicated to what degree this sound corresponded to a “B.”

The experiment was controlled with Nijmegen Experiment Set-Up (NESU) software, and stimuli were presented binaurally, one at a time, over Sennheiser closed headphones at a comfortable listening level. Participants responded by pressing one of four response buttons. Stimuli were presented with 3000 ms intervals between onsets.

## B. Results and discussion

Analyses of variance (ANOVAs) were done on averaged goodness ratings, with vowel duration and fricative voicing as within-subjects factors and native language as between-subjects factor. As a measure of effect size, *partial*  $\eta^2$  is reported. A *partial*  $\eta^2$  of 0.01 is considered to indicate a small effect size, a *partial*  $\eta^2$  of 0.06 a medium, and of 0.14 a large effect size. Mean responses are presented in Table I.

First, for the “V” target, Fig. 1(a) shows that both for the Dutch and for the English listeners, as expected, goodness ratings are higher for items with a long vowel than for items with a short vowel [vowel duration:  $F(1,30)=209.23, p < 0.001, \textit{partial} \eta^2=0.88$ ] and higher for items with more voiced fricatives than for items with less voiced fricatives [fricative voicing:  $F(10,300)=8.86, p < 0.001, \textit{partial} \eta^2=0.23$ ]. Crucially, however, the use of vowel duration differs for the Dutch and English listeners in part of the voicing continuum, as reflected in a significant interaction among vowel duration, fricative voicing, and native language [ $F(10,300)=2.79, p < 0.01, \textit{partial} \eta^2=0.09$ ]. For the English listeners, in the condition with the short preceding vowel, goodness ratings are higher for items with more voiced fricatives than for items with less voiced fricatives, as expected [ $F(10,150)=9.36, p < 0.001, \textit{partial} \eta^2=0.38$ ], but in the condition with the long preceding vowel, there is no significant effect of fricative voicing [ $F(10,150)=1.49, p > 0.1$ ]. Thus, English listeners give stimuli with more voiced fricatives and stimuli with less voiced fricatives similar goodness ratings for “V” when they are preceded by a long vowel. For those listeners, indeed, there is an interaction between vowel duration and fricative voicing [ $F(10,150)=4.97, p < 0.001, \textit{partial} \eta^2=0.25$ ]. Thus, long vowel durations affect the English listeners’ goodness ratings more than short vowel durations do. For the Dutch listeners, on the other hand, there is no interaction between vowel duration and fricative voicing [ $F(10,150) < 1$ ]. For those listeners, goodness ratings are higher for items with more voiced fricatives than for items with less voiced fricatives [ $F(10,150)=2.88, p < 0.01, \textit{partial} \eta^2=0.16$ ] for both vowel durations alike, and goodness ratings are higher for items with a long vowel than for items with a short vowel [ $F(1,15)=91.70, p < 0.001, \textit{partial} \eta^2=0.86$ ]. This pattern is in line with the results reported in Broersma (2005, 2008) that for English listeners, when vowel duration and fricative voicing do not match, the long vowel but not the short vowel outweighs fricative voicing, but for Dutch listeners, long and short vowels affect /v-f/ categorization to the same extent.

Second, for the “F” target, Fig. 1(b) shows that for Dutch and English listeners, as expected, goodness ratings are higher for items with a short vowel than for items with a long vowel [ $F(1,30)=27.13, p < 0.001, \textit{partial} \eta^2=0.48$ ] and higher for items with less voiced fricatives than for items with more voiced fricatives [ $F(10,300)=14.69, p < 0.001, \textit{partial} \eta^2=0.33$ ]. Crucially, vowel duration affects Dutch listeners’ goodness ratings for “F” less than the English listeners’ ratings, as reflected in an interaction between vowel duration and native language [ $F(1,30)=4.50, p < 0.05, \textit{partial} \eta^2=0.13$ ]. For the English listeners, the effect of vowel

duration, in the expected direction, is significant [ $F(1,15)=103.03, p < 0.001, \textit{partial} \eta^2=0.87$ ]; for the Dutch listeners, it is not [ $F(1,15)=2.74, p > 0.1$ ]. For the condition with the short preceding vowel, there is no difference between the Dutch and English’ listeners’ goodness ratings, [ $F(1,30) < 1$ ], but for the condition with the long preceding vowel, Dutch listeners give higher scores than English listeners do [ $F(1,30)=11.01, p < 0.01, \textit{partial} \eta^2=0.27$ ]. Thus, not all vowel durations but specifically long vowel durations lead to different ratings for English listeners than for Dutch listeners. This is in line with the finding for the “V” target that long vowels affect the English listeners but not the Dutch listeners particularly strongly.

Comparing “V” and “F” targets, in the condition with the long vowel, Dutch listeners give high goodness ratings for the “V” target, but also relatively high goodness ratings for the “F” target. Thus, whereas the English listeners perceive stimuli with a long vowel as a good “V” and as a poor “F,” the Dutch listeners perceive them as good exemplars of “V” but also as relatively good exemplars of “F”, showing again that vowel duration is not as decisive for the Dutch listeners’ perception of the goodness of the voicing of those fricatives as it is for the English listeners’ perception.

For the “Z” target [Fig. 2(a)] and for the “S” target [Fig. 2(b)], crucially, vowel duration significantly affects goodness ratings too. For the “Z” target, items with a long vowel receive higher goodness ratings than items with a short vowel [ $F(1,30)=68.43, p < 0.001, \textit{partial} \eta^2=0.70$ ], and for the “S” target, items with a short vowel receive higher goodness ratings than items with a long vowel [ $F(1,30)=50.40, p < 0.001, \textit{partial} \eta^2=0.63$ ] for Dutch and English listeners alike. Further, the effect of fricative voicing is also as expected: for the “Z” target, goodness ratings are higher for items with more voiced fricatives than for items with less voiced fricatives [ $F(10,300)=31.03, p < 0.001, \textit{partial} \eta^2=0.51$ ], and for the “S” target, goodness ratings are higher for items with less voiced fricatives than for items with more voiced fricatives [ $F(10,300)=32.36, p < 0.001, \textit{partial} \eta^2=0.52$ ] for Dutch and English listeners alike.

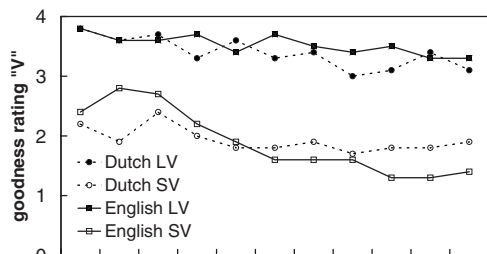
Finally, comparing the /v-f/ and the /z-s/ contrast, there is an interaction between vowel duration and contrast [ $F(1,30)=4.39, p < 0.05, \textit{partial} \eta^2=0.13$ ], confirming that the effect of vowel duration is larger for the /v-f/ contrast than for the /z-s/ contrast.

In summary, for the /v-f/ and the /z-s/ contrast, both vowel duration and fricative voicing affect the goodness ratings as expected: long vowels lead to higher goodness ratings than short vowels for voiced targets and vice versa for voiceless targets, and more voiced fricatives lead to higher goodness ratings than more voiceless fricatives for voiced targets and vice versa for voiceless targets. Effect sizes are large both for vowel duration and fricative voicing, but larger for the former than for the latter. Importantly, the effect of vowel duration is found not only for the /v-f/ contrast but also for the /z-s/ contrast.

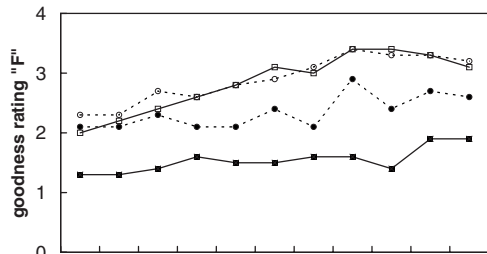
For the /v-f/ contrast, vowel duration and, in particular, long vowels affect English listeners’ goodness ratings more than Dutch listeners’ ratings. These results are in line with the phonetic categorization results from Broersma (2005,

TABLE I. Results Experiments 1 and 2.

		Mean goodness ratings				
		Dutch participants		English participants		
	Step fricative voicing continuum	Long vowel	Short vowel	Long vowel	Short vowel	
Experiment 1	“V” target	1(voiced)	3.8	2.2	3.8	2.4
		2	3.6	1.9	3.6	2.8
		3	3.7	2.4	3.6	2.7
		4	3.3	2.0	3.7	2.2
		5	3.6	1.8	3.4	1.9
		6	3.3	1.8	3.7	1.6
		7	3.4	1.9	3.5	1.6
		8	3.0	1.7	3.4	1.6
		9	3.1	1.8	3.5	1.3
		10	3.4	1.8	3.3	1.3
		11	3.1	1.9	3.3	1.4
“F” target	1(voiced)	2	2.1	2.3	1.3	2.0
		3	2.3	2.7	1.4	2.4
		4	2.1	2.6	1.6	2.6
		5	2.1	2.8	1.5	2.8
		6	2.4	2.9	1.5	3.1
		7	2.1	3.1	1.6	3.0
		8	2.9	3.4	1.6	3.4
		9	2.4	3.3	1.4	3.4
		10	2.7	3.3	1.9	3.3
		11	2.6	3.2	1.9	3.1
		“Z” target	1(voiced)	2	3.9	3.4
3	3.6			2.4	3.4	2.6
4	3.4			2.6	3.4	2.5
5	3.4			2.6	3.3	2.1
6	2.9			2.6	3.2	2.2
7	3.3			1.8	3.0	1.5
8	3.2			1.8	2.8	1.6
9	2.4			1.3	2.8	1.4
10	2.3			1.4	2.3	1.2
11	2.1			1.3	2.4	1.3
11	2.2			1.4	2.4	1.3
“S” target	1(voiced)	2	1.3	1.8	2.0	2.1
		3	1.3	2.0	1.4	2.6
		4	1.5	2.1	1.7	2.2
		5	1.4	2.1	1.8	2.4
		6	1.9	2.6	2.1	2.7
		7	1.9	2.8	2.1	2.9
		8	2.1	3.3	2.3	3.4
		9	1.9	3.3	2.3	3.4
		10	2.3	3.4	2.6	3.8
		11	2.6	3.9	2.9	3.7
		11	2.6	3.9	2.4	3.6
		Mean percentage “voiced” responses				
		Dutch participants		English participants		
	Step fricative voicing continuum	Long vowel	Short vowel	Long vowel	Short vowel	
Experiment 2	/v-f/ contrast	1(voiced)	93.4	36.9	91.6	32.8
		2	94.7	34.1	93.8	30.6
		3	92.2	31.9	95.9	25.0
		4	93.1	23.1	95.6	18.1
		5	90.6	15.6	95.0	11.9
		6	88.1	12.5	94.4	6.3
		7	83.1	8.4	88.1	8.4
		8	77.2	5.3	87.8	6.9
		9	69.1	6.3	87.2	8.4
		10	67.8	4.7	88.8	6.6
		11	69.1	5.0	85.3	5.6
/z-s/ contrast	1(voiced)	2	92.5	58.4	92.5	38.4
		3	95.0	55.3	90.0	31.3
		4	94.4	52.8	88.8	31.6
		5	90.9	51.9	88.1	29.7
		6	88.4	37.2	86.6	24.4
		7	84.4	22.8	79.7	19.7
		8	74.1	13.4	78.4	10.6
		9	65.3	8.1	66.6	6.6
		10	61.6	5.9	65.0	5.9
		11	57.2	5.3	62.2	5.0
		11	56.6	2.5	62.5	5.3



(a) /v/ 3 5 7 9 /f/



(b) /v/ 3 5 7 9 /f/

FIG. 1. Mean goodness rating as a function of the place on an 11-step /v/ to /f/ continuum, preceding vowel duration (LV: long vowel; SV: short vowel), and participants' native language: (a) target "V" and (b) target "F."

2008) that showed a larger effect of vowel duration for English listeners than for Dutch listeners and an especially large effect of long vowel durations when combined with voiceless fricatives for English listeners. The present results show that those findings were not an artifact of the design of Broersma (2005, 2008), where vowel duration was kept constant for each participant, but is also found when vowel duration is varied for each participant.

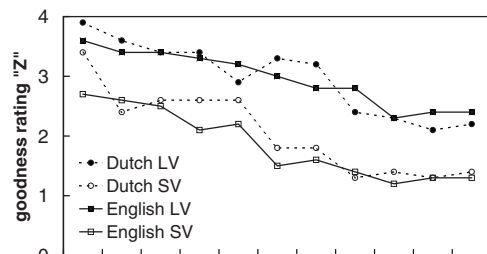
Experiment 2 further investigates the use of vowel duration when it varies within subjects with a phonetic categorization task. The experiment is identical to that in Broersma (2005, 2008), except that all participants hear both the phonetically long vowel and the phonetically short vowel. Stimuli thus vary in vowel duration and in fricative voicing for each participant. Cue trading between temporal and spectral cues has often been found (Fitch *et al.*, 1980; Sinnott and Saporita, 2000), and a trading relation between vowel duration and fricative voicing is anticipated here too. Note that whereas fricative voicing varies in 11 steps, vowel duration varies in only two steps, which might make the latter variable relatively salient (Bohn, 1995). The experiment addresses the same three questions; first, whether there is still a difference between the Dutch and the English listeners' use of vowel duration for the /v-f/ contrast when it varies within subjects, second, whether vowel duration is used as a perceptual cue for the /z-s/ contrast, like in Experiment 1, but unlike in Broersma (2005, 2008), and third, whether Dutch and English listeners use vowel duration differentially for the /z-s/ contrast.

### III. EXPERIMENT 2

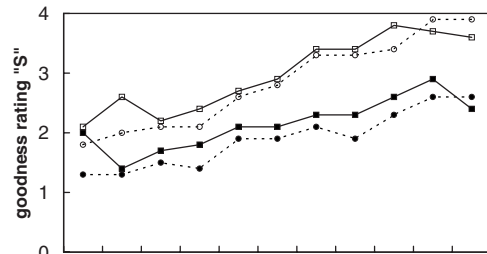
#### A. Method

##### 1. Participants

Participants were 16 native speakers of Dutch and 16 native speakers of British English, matching the description given for Experiment 1, none of whom had participated in Experiment 1.



(a) /z/ 3 5 7 9 /s/



(b) /z/ 3 5 7 9 /s/

FIG. 2. Mean goodness rating as a function of the place on an 11-step /z/ to /s/ continuum, preceding vowel duration (LV: long vowel; SV: short vowel), and participants' native language: (a) target "Z" and (b) target "S."

## 2. Materials

The same materials were used as in Broersma (2005) and in Experiment 1.

## 3. Design

As in Experiment 1, stimuli were blocked by contrast but, unlike in Broersma (2005), not by vowel duration. Thus, each participant heard both the phonetically long and the phonetically short vowel within the same experiment and within the same block. As the carrier for the /z-s/ contrast contained an /f/, the /v-f/ contrast was always tested before the /z-s/ contrast. Each block contained 20 repetitions of all combinations of 11 fricatives and two carriers, yielding 440 trials per block. Items were semi-randomized, such that the same step of the fricative continuum could not occur twice in succession. There was no practice with the crucial contrasts, only with the /b-p/ contrast (as in Experiment 1).

## 4. Procedure

Participants were tested one at a time in a quiet room. They received written instructions in their native language that they would hear a series of English nonwords, which would be similar except for the final sound. They were instructed to decide whether this was a "V" or an "F," or a "Z" or an "S," respectively, and to indicate their response by pressing one of two buttons, labeled "V" and "F," or "Z" and "S." The experiment started with a short practice during which the participants categorized final stops as "B" or "P."

The experiment was controlled with NESU experimental software. Stimuli were presented binaurally, one at a time, over Sennheiser closed headphones at a comfortable listening level. Participants responded by pressing one of two response buttons. No time limit was imposed for the responses. After each button press, presentation of the next item started.

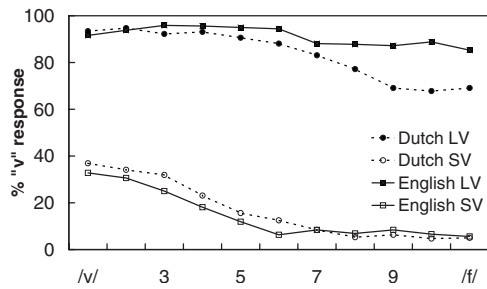


FIG. 3. Mean percentage of “V” responses as a function of the place on an 11-step /v/ to /f/ continuum, preceding vowel duration (LV: long vowel; SV: short vowel), and participants’ native language.

## B. Results and discussion

Two responses with reaction times longer than 10 000 ms were removed. As Figs. 3 and 4 show, for both contrasts, the effect of vowel duration is very large. The categorization curves are very shallow, and the mean percentages of “voiced” responses for the conditions with the long and short preceding vowels do not overlap, except for the Dutch listeners’ responses for the /z-s/ contrast, where they just touch (with 57% “Z” responses for step 11 in the condition with the long preceding vowel and 58% for step 1 in the condition with the short preceding vowel). ANOVAs were done on arcsine transformed proportions of “voiced” responses, with vowel duration and fricative voicing as within-subjects factors and native language as between-subjects factor. Mean responses are presented in Table I.

Figure 3 shows that especially for the /v-f/ contrast, the effect of vowel duration is extremely large and the effect of fricative voicing is relatively small, both for the Dutch and for the English listeners. Indeed, stimuli with long vowels receive more “V” responses than stimuli with short vowels, both for the Dutch listeners [ $F(1, 15)=59.43, p<0.001, \text{partial } \eta^2=0.80$ ] and for the English listeners [ $F(1, 15)=160.95, p<0.001, \text{partial } \eta^2=0.92$ ]. Items with more voiced fricatives also receive significantly more “V” responses than items with less voiced fricatives for the Dutch listeners [ $F(10, 150)=12.83, p<0.001, \text{partial } \eta^2=0.46$ ] and for the English listeners [ $F(10, 150)=6.76, p<0.001, \text{partial } \eta^2=0.31$ ].

Crucially, however, there are differences between the Dutch and English listeners’ categorization responses, reflected in a significant three-way interaction among vowel

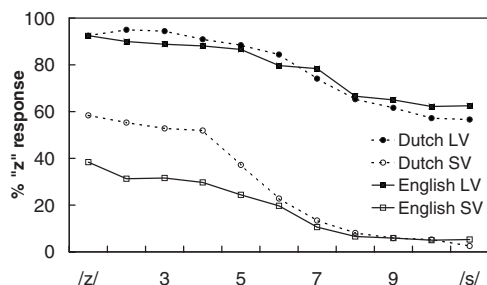


FIG. 4. Mean percentage of “Z” responses as a function of the place on an 11-step /z/ to /s/ continuum, preceding vowel duration (LV: long vowel; SV: short vowel), and participants’ native language.

duration, fricative voicing, and native language [ $F(10, 300)=1.90, p<0.05, \text{partial } \eta^2=0.06$ ]. As Fig. 3 shows, in the condition with the long preceding vowel, the English listeners give more “V” responses than the Dutch listeners to items with less voiced fricatives. Indeed, in the condition with the long preceding vowel, the interaction between fricative voicing and native language is significant [ $F(10, 300)=2.60, p<0.005, \text{partial } \eta^2=0.08$ ]. For the condition with the short preceding vowel, there is no interaction between fricative voicing and native language [ $F(10, 300)<1$ ].

For the /z-s/ contrast, Fig. 4 again shows a large effect of vowel duration and a relatively small effect of fricative voicing, both for the Dutch and for the English listeners. Similar to the /v-f/ contrast, stimuli with long vowels receive more “Z” responses than stimuli with short vowels, both for the Dutch listeners [ $F(1, 15)=33.87, p<0.001, \text{partial } \eta^2=0.69$ ] and for the English listeners [ $F(1, 15)=61.49, p<0.001, \text{partial } \eta^2=0.80$ ]. Items with more voiced fricatives also receive significantly more “Z” responses than items with less voiced fricatives for the Dutch listeners [ $F(10, 150)=20.58, p<0.001, \text{partial } \eta^2=0.58$ ] and for the English listeners [ $F(10, 150)=14.21, p<0.001, \text{partial } \eta^2=0.49$ ].

Also similar to the /v-f/ contrast, there is a difference between the Dutch and English listeners’ categorization responses, reflected in a significant three-way interaction among vowel duration, fricative voicing, and native language [ $F(10, 300)=2.09, p<0.05, \text{partial } \eta^2=0.07$ ]. Figure 4 shows that the difference between the Dutch and English listeners’ responses is now located on the other side of the graph than for the /v-f/ contrast: in the condition with the short preceding vowel the English listeners give fewer “Z” responses than the Dutch listeners to items with more voiced fricatives. In that condition, the interaction between fricative voicing and native language is significant [ $F(10, 300)=2.98, p<0.001, \text{partial } \eta^2=0.09$ ]. For the condition with the long preceding vowel, there is no interaction between fricative voicing and native language [ $F(10, 300)<1$ ].

The results for the /v-f/ and /z-s/ contrasts are thus to a large extent similar: both Dutch and English listeners use vowel duration and fricative voicing as a perceptual cue, effect sizes are large both for vowel duration and fricative voicing but larger for the former than for the latter, and for both contrasts, differences between the Dutch and English listeners’ responses appear where the stimuli contain mismatching cues. However, where these differences occur is different for both contrasts. For the /v-f/ contrast, English listeners give more “voiced” responses than the Dutch listeners do to items with less voiced fricatives preceded by a long vowel, and for the /z-s/ contrast, English listeners give more “voiceless” responses than the Dutch listeners do to items with more voiced fricatives preceded by a short vowel. Thus, in these places, the English listeners rely more on vowel duration than the Dutch listeners do.<sup>2</sup>

The results for the /v-f/ contrast are in line with those from Broersma (2005), where English listeners also gave more “voiced” responses than Dutch listeners did to long vowels combined with less voiced fricatives. For the /z-s/ contrast, on the other hand, there was no effect of vowel duration in that study. In the present study, the effect of

vowel duration is also smaller and the effect of fricative voicing is larger for the /z-s/ contrast than for the /v-f/ contrast (an ANOVA with contrast as an additional within-subjects factor shows significant interactions between vowel duration and contrast [ $F(1,30)=11.37, p<0.01, \text{partial } \eta^2=0.28$ ] and between fricative voicing and contrast [ $F(10,300)=5.90, p<0.001, \text{partial } \eta^2=0.16$ ]). However, the effect size of vowel duration for the /z-s/ contrast is still very large, and crucially, vowel duration affects Dutch and English listeners differentially.

#### IV. GENERAL DISCUSSION

Three questions were raised in the Introduction. First, would Dutch and English listeners use vowel duration differently for the /v-f/ contrast when it was varied for each participant? Both experiments showed that this was indeed the case. Thus, the finding of Broersma (2005, 2008) that Dutch listeners used vowel duration less than English listeners did was not an artifact of the between-subjects manipulation of vowel duration in that study. In the present experiments, where the design did not discourage the use of vowel duration as a perceptual cue, Dutch listeners still used it less than English listeners did.

The second question was whether there would be an effect of vowel duration for the /z-s/ contrast when vowel duration varied within subjects. Indeed, in both experiments, a clear effect of vowel duration was found. Thus, the lack of such an effect in Broersma (2005) was most likely due to the design of that study. The finding that the English listeners used vowel duration for the final /z-s/ distinction is in line with earlier research (e.g., Derr and Massaro, 1980; Hogan and Rozsypal, 1980; Raphael, 1972). The effect of vowel duration was smaller for the /z-s/ contrast than for the /v-f/ contrast in both experiments.

The third question was whether Dutch and English listeners would use vowel duration differently for the /z-s/ contrast too. In Experiment 1, there were no differences between the Dutch and English listeners' use of vowel duration for the /z-s/ contrast. Experiment 2 (involving a more sensitive paradigm) showed, however, that English listeners used vowel duration more than Dutch listeners did. Thus, the differences between Dutch and English listeners' use of vowel duration as a perceptual cue are not limited to the /v-f/ contrast, but occur for the /z-s/ contrast too.

Experiment 2 was a replication of Broersma (2005, 2008), except for the within-subjects manipulation of vowel duration. The results show that this design choice affected the listeners' use of vowel duration as expected. In Experiment 2, a large effect size of vowel duration was found for both contrasts. (This was also the case in Experiment 1, where vowel duration was also varied for each participant.) In Broersma (2005, 2008), on the other hand, where vowel duration was varied only between subjects, the effect size of vowel duration for the /v-f/ contrast was medium (with a *partial*  $\eta^2$  of 0.12), and there was no significant effect of vowel duration for the /z-s/ contrast.

What do these results mean in terms of perceptual accuracy? The differences between Dutch and English listeners,

although very robust, were also especially prominent where perceptual cues mismatched. Further, Dutch listeners did use vowel duration as a final voicing cue, and the differences between the Dutch and the English listeners' use of vowel duration were relatively small compared to the overall size of the effect of vowel duration. This may explain why in a previous study (Broersma, 2005, Experiment 1) Dutch listeners were found to categorize unedited tokens of final voiced and voiceless fricatives as accurately as English listeners did and as accurately as word-initial fricatives. Their use of vowel duration might be sufficiently adequate to recognize final voiced and voiceless fricatives under normal circumstances (i.e., when perceptual cues do not mismatch). Further, they might successfully use other temporal, spectral, and power properties of the stimuli (e.g., Jongman *et al.*, 2000; Silbert and De Jong, 2008) to distinguish between voiced and voiceless final fricatives. Therefore, in normal language use, the differences in cue weighting demonstrated here might not hinder Dutch listeners' comprehension of English speech much, in line with the predictions of the PAM (Best *et al.*, 1988; Best and Tyler, 2007).

Finally, what do these results mean for the transfer of the use of perceptual cues from the L1 to the L2? It was argued in the Introduction that Dutch listeners were well equipped for the use of vowel duration as a perceptual cue for final fricative voicing. Dutch provides experience with the use of vowel duration as a perceptual cue for vowel length contrasts and even for intervocalic obstruent voicing contrasts. Despite their native language experience with the use of vowel duration as a perceptual cue, Dutch listeners used it less than English listeners did for the English final fricative voicing contrasts. Thus, the Dutch listeners (advanced L2 learners with ample exposure to English) did not seem to apply their L1 experience with vowel duration when listening to English contrasts. Whether their L1 experience affected the use of vowel duration for final fricative voicing at all, the present data cannot reveal, but the results showed that the Dutch listeners did not rely on this cue as strongly as native English listeners did. The findings extend those from Broersma (2005, 2008) and seem to reflect a robust difference between Dutch and English listeners' use of vowel duration that is not limited to a single phoneme contrast or to a particular experimental design or paradigm. Thus, experience with a perceptual cue for a contrast in the L1 may not suffice for the efficient use of this cue for a different contrast in the L2, or even for a similar contrast in a different phonetic position in the L2.

#### ACKNOWLEDGMENTS

This research was supported by a Veni grant from the Netherlands Organisation for Scientific Research (NWO). Many thanks to Alan Garnham of the Laboratory of Experimental Psychology, University of Sussex, and to Anne Cutler of the Max Planck Institute for Psycholinguistics, Nijmegen, for the opportunity to test participants in their laboratories. I am grateful to Marco van de Ven for testing Dutch participants and to an anonymous reviewer for helpful comments.



- <sup>1</sup>Voiced and voiceless fricatives of equal duration were needed in order to create the continua. As frication noise is generally longer for voiceless final fricatives than for voiced final fricatives (Crystal and House, 1988; Silbert and De Jong, 2008) and listeners use frication duration as a perceptual cue for final fricative voicing (Watson, 1983), shortening the voiceless fricatives may have made them perceptually more voiced.
- <sup>2</sup>In Broersma (2008), English listeners' use of vowel duration changed during the experiment. Here, this was not the case. The data for each contrast were split into seven parts. In an ANOVA with part as an additional within-subjects factor, there were no interactions with or main effect of part, showing that the use of vowel duration or fricative voicing did not change during the experiment.
- Best, C. T., McRoberts, G. W., and Sithole, N. M. (1988). "Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants," *J. Exp. Psychol. Hum. Percept. Perform.* **14**, 345–360.
- Best, C. T., and Strange, W. (1992). "Effects of phonological and phonetic factors on cross-language perception of approximants," *J. Phonetics* **20**, 305–330.
- Best, C. T., and Tyler, M. D. (2007). "Nonnative and second-language speech perception: Commonalities and complementarities," in *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, edited by O.-S. Bohn and M. J. Munro (John Benjamins, Amsterdam), pp. 13–34.
- Bohn, O.-S. (1995). "Cross-language speech perception in adults: First language transfer doesn't tell it all," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York Press, Baltimore, MD), pp. 279–303.
- Bohn, O.-S. and Munro, M. J., eds. (2007). *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege* (John Benjamins, Amsterdam).
- Booij, G. (1995). *The Phonology of Dutch* (Oxford University Press, Oxford).
- Bradlow, A. R. (1996). "A perceptual comparison of the /i/-/e/ and /u/-/o/ contrasts in English and Spanish: Universal and language-specific aspects," *Phonetica* **53**, 55–85.
- Broersma, M. (2005). "Perception of familiar contrasts in unfamiliar positions," *J. Acoust. Soc. Am.* **117**, 3890–3901.
- Broersma, M. (2008). "Flexible cue use in nonnative phonetic categorization," *J. Acoust. Soc. Am.* **124**, 712–715.
- Crowther, C. S., and Mann, V. (1992). "Native language factors affecting use of vocalic cues to final consonant voicing in English," *J. Acoust. Soc. Am.* **92**, 711–722.
- Crystal, T. H., and House, A. S. (1988). "A note on the durations of fricatives in American English," *J. Acoust. Soc. Am.* **84**, 1932–1935.
- Denes, P. (1955). "Effect of duration on the perception of voicing," *J. Acoust. Soc. Am.* **27**, 761–764.
- Derr, M. A., and Massaro, D. W. (1980). "The contribution of vowel duration, F0 contour, and frication duration as cues to the /juz/-/jus/ distinction," *Percept. Psychophys.* **27**, 51–59.
- Fitch, H. L., Halwes, T., Erickson, D. M., and Liberman, A. M. (1980). "Perceptual equivalence of two acoustic cues for stop-consonant manner," *Percept. Psychophys.* **27**, 343–350.
- Flege, J. E. (1995). "Second language learning: Theory, findings, and problems," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York Press, Baltimore, MD), pp. 233–272.
- Flege, J. E., and Hillenbrand, J. (1986). "Differential use of temporal cues to the /s/-/z/ contrast by native and non-native speakers of English," *J. Acoust. Soc. Am.* **79**, 508–517.
- Hogan, J. T., and Rozsypal, A. J. (1980). "Evaluation of vowel duration as a cue to the voicing distinction in the following word-final consonant," *J. Acoust. Soc. Am.* **67**, 1764–1771.
- Jones, D. (1950). *The Phoneme: Its Nature and Use* (W. Heffer & Sons, Cambridge).
- 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.
- Peterson, G. E., and Lehiste, I. (1960). "Duration of syllable nuclei in English," *J. Acoust. Soc. Am.* **32**, 693–703.
- 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.
- Silbert, N., and De Jong, K. (2008). "Focus, prosodic context, and phonological feature specification: Patterns of variation in fricative production," *J. Acoust. Soc. Am.* **123**, 2769–2779.
- Sinnott, J. M., and Saporita, T. A. (2000). "Differences in American English, Spanish, and monkey perception of the say-stay trading relation," *Percept. Psychophys.* **62**, 1312–1319.
- Slis, I. H., and Cohen, A. (1969a). "On the complex regulating the voiced-voiceless distinction I," *Lang Speech* **12**, 80–102.
- Slis, I. H., and Cohen, A. (1969b). "On the complex regulating the voiced-voiceless distinction II," *Lang Speech* **12**, 137–155.
- Slis, I. H., and Van Heugten, M. (1989). "Voiced-voiceless distinction in Dutch fricatives," in *Linguistics in The Netherlands 1989*, edited by H. Bennis and A. Van Kemenade (Foris, Dordrecht, The Netherlands), pp. 123–132.
- Stevenson, D. C. (1979). "Categorical perception and selective adaptation phenomena in speech," Ph.D. dissertation, University of Alberta, Edmonton, Canada.
- Strange, W. (1995). *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (York Press, Baltimore, MD).
- Van den Berg, R. J. H. (1987). "Effects of durational factors on the perception of voicing in Dutch two-obstruent sequences," *J. Phonetics* **15**, 259–271.
- Van den Berg, R. J. H. (1988). "The perception of voicing in Dutch two-obstruent sequences: A comparison of synthetic and natural speech," *J. Phonetics* **16**, 171–180.
- Van den Berg, R. J. H. (1989). "Perception of voicing in Dutch two-obstruent sequences: Covariation of voicing cues," *Speech Commun.* **8**, 17–25.
- Watson, I. (1983). "Cues to the voicing contrast: A survey," *Cambridge Papers in Phonetics and Experimental Linguistics* **2**, 1–34.