

## Final devoicing: Production and perception studies\*

Scott Myers

University of Texas at Austin

Final devoicing is a pattern of phonological distribution in which both voiced and voiceless obstruents occur in a language, but at the end of a particular prosodic domain (Selkirk 1978, 1986) only voiceless obstruents occur.<sup>1, 2, 3</sup> There are examples from all over the world, involving both phonological word and syllable domains (cf. Passy 1891: 160; Grammont 1933: 365; Locke 1983: 118; Lombardi 1995, 1999; Blevins 2006; Harris 2009):

---

\* I would like to thank the following people for their helpful comments on this work: Juliette Blevins, Shigeto Kawahara, two anonymous reviewers, and audiences at the University of Massachusetts at Amherst and the Acoustical Society of America meeting in Portland. Thanks also to Lisa Selkirk, who I was lucky to have as my PhD supervisor, for her mentorship, her friendship, and the example she has always provided of elegant and insightful inquiry into how language works.

<sup>1</sup> By "voiced" obstruents I mean obstruents that are distinguished from "voiceless" ones by having a higher proportion of vocal fold pulsing and a lower proportion of aperiodic noise. They are generally distinguished by other phonetic properties as well (Lisker 1986, Kingston and Diehl 1994), and in some of these cases periodic pulsing may not be the primary cue (Jessen and Ringen 2002).

<sup>2</sup> Descriptions of the final devoicing pattern are often frustratingly vague. The position of neutralization is sometimes simply described as "final", without making clear what domain it is final in (e.g. Dambriunas et al. 1966: 17), and it is often not clear whether the author checked other possible domains (e.g. whether devoicing occurs at the end of a word if the word is non-final in the phrase). Most of the descriptions cited here are also based on transcriptions, so they are inherently vague as to whether the devoicing effect is gradient or categorical.

<sup>3</sup> The neutralization of the voicing contrast in final position in some languages is incomplete, i.e. there are measurable and perceptible differences between alternating and nonalternating final voiceless consonants (Dinnsen and Charles-Luce 1984, Port and O'Dell 1985, Slowiaczek and Dinnsen 1985, Charles-Luce and Dinnsen 1987, Slowiaczek and Szymanska 1987, Warner et al. 2004, and Dmitrieva 2005). Ernestus and Baayen (2006: 47) suggest that words with alternating final voiceless obstruents are influenced in production by the activation of the corresponding voiced-final items in their paradigm.

(1) (a) All word-final obstruents are voiceless.

- Slavic: Russian (Padgett 2002), Czech (Heim 1976: 14), Slovak (Rubach 1993: 283), Bulgarian (Scatton 1984: 20), Polish (Rubach 1984: 206)
- Romance: Walloon (Francard and Morin 1986: 457), Friulian (Baroni and Vanelli 2000: 27), Old French (Ewert 1933: 75, 97), Ferrarese Italian (stops only - Dinnsen and Eckman 1978: 5)
- Germanic: Dutch (Booij 1995: 22), German (Jessen and Ringen 2002), Gothic (fricatives only - Wright 1899: 62-67; Hock 1991: 43), Old English (fricatives only - Hock 1991: 43)
- Saranda Ekklisies Greek (stops only - Newton 1972: 103)
- Sanskrit (Whitney 1879: 46)
- Daragözu Arabic and Maltese Arabic (Abu Mansour 1996: 213, 215)
- Uyghur (Hahn 1991: 84)
- Fur (Jakobi 1990: 35)
- Luo (Tucker 1994: 35)
- Afar (Bliese 1981: 242)
- Basque (Hualde 1991: 13)

(b) All syllable-final obstruents are voiceless.

- Takelma (Sapir 1990: 35)
- Wintu (Pitkin 1984: 26)
- West Tarangan (Nivens 1992: 147)
- Romansch (Montreuil 1999: 531)
- Catalan (Hualde 1992: 393)
- Breton (Krämer 2000: 641)
- Haisla (Bach 1996: 5)
- Ron (Jungraithmayr 1970: 21)
- Malay (Ahmad 2005: 55)
- Turkish (stops only - Clements and Keyser 1983: 59-60)
- Buriat (Poppe 1960: 10)

- Efik (Cook 1969: 36)
- Manipuri (Singh 2000: 13-16)
- Thai (Iwasaki and Ingkaphirom 2005: 4)
- Vietnamese (Thompson 1965: 23)
- Various Sino-Tibetan languages (Thurgood and LaPolla 2003)

An optional pattern of word-final obstruent devoicing has been reported as a distinguishing characteristic of a number of English dialects, e.g. those of some African American communities (Wolfram 1969: 51, Luelsdorf 1975: 42), the Appalachian region (Wolfram and Christian 1976: 63), or the Maori people in New Zealand (Holmes 1996).

The pattern is also reflected in language acquisition. Already in prelinguistic babbling voiceless consonants greatly outnumber voiced consonants in utterance-final position (Oller et al. 1976). When children begin to produce words, they often have a stage where they produce both voiced and voiceless obstruents, but in word-final position only voiceless ones (Velten 1943: 283, Smith 1973: 37, Smith 1979: 22, Flege 1982). Persistence of this error type later into the acquisition process is recognized as a speech disorder (Hodson and Paden 1981: 371, Cutts and Jensen 1983, Ingram 1989: 115). Systematic devoicing of final voiced target consonants is also a prevalent error type in second language acquisition, even in cases where final devoicing is not a characteristic of either the target language or the learner's native language (Eckman 1981, Flege and Davidian 1984, Flege, Munro and Skelton 1992, Major and Audree 1996, Broselow, Chen, and Wang 1998).

Scholars have long related this phonological pattern of final devoicing to the phonetics of prepausal position (e.g. Sievers 1901: 289-290; Jespersen 1926: 101; Bloomfield 1933: 373; Lindblom 1983: 237). There is no vocal fold vibration during pause, so devoicing in prepausal position can be seen as assimilation to this voiceless state (Lightner 1972: 332-333, Ingram 1989: 35). Non-speech breathing is characterized by a wide glottal aperture to facilitate air passage, and speakers begin spreading the vocal folds in anticipation before they are done producing speech (Sweet 1877: 65; Lisker et al. 1969: 1545; Klatt and Klatt 1990, Shadle 1997: 42; Jessen 1998; Slifka 2006). In addition to this coarticulatory effect of the transition from speech to nonspeech, voicing in

utterance-final position is also hampered by a decline in subglottal pressure over the course of the utterance (Westbury and Keating 1986: 156). The result is a gradual breakdown in voicing as one approaches pause, often passing through nonmodal voicing before a final voiceless interval. Such utterance-final devoicing has been found in instrumental acoustic studies in English (Haggard 1978, Docherty 1992, Smith 1997), French (Smith 1999, 2003), Finnish (Lehtonen 1970: 45, Myers and Hansen 2007), and Kinyarwanda (Myers 2005), and noted as well in many transcription-based studies (e.g. Michelsen's 1988 study of the Lake Iroquoian languages). Oller and Smith (1977) found utterance-final vowel devoicing to be a regular feature of prelinguistic babbling.

I would propose that this coarticulatory utterance-final devoicing is the initial impetus for a sound change that results in phonological word-final devoicing. The first step of this transformation would be that utterance-final devoicing affects the perception of voicing contrasts in utterance-final position, inducing a tendency among listeners to identify utterance-final obstruents as voiceless. This would be expected since utterance-final devoicing diminishes voicing during the constriction period, a demonstrated perceptual cue for voicing (Raphael 1971, Wolf 1978, Smith 1979, Hogan and Roszypal 1980, Kingston and Diehl 1994). As Blevins (2006) points out, lengthening of the utterance-final consonant (Lindblom 1968) could have the same perceptual effect, since listeners are also sensitive to the fact that voiceless obstruents are longer than corresponding voiced ones (Denes 1955). Listeners are generally adept at compensating in perception for coarticulatory effects (Lindblom and Studdert-Kennedy 1967, Mann and Repp 1980), but they can fail to do so, leading to hypocorrection (Ohala 1981, 1993). In this case, the result of failing to completely compensate for the devoicing effect of utterance-final position would be a tendency to identify utterance-final obstruents as voiceless.

The second step of the sound change would be that the listener generalizes over the voicing categories he or she has identified in this way, and concludes that obstruents in this position are voiceless. This generalization is a phonological restriction on category distribution. Pierrehumbert (2001: 152) has shown how in an exemplar-based model even a small bias in identification of this sort can over time lead to such a neutralization in contrast between two speech sound categories. The extension of the pattern from

utterance-final to word- and syllable-final positions would be an analogical extension based on the fact that every utterance-final consonant is also word- and syllable-final (Ewert 1933: 75; Westbury and Keating 1986: 161; Hock 1991: 239).

The third step of the sound change is the spread of the pattern from the individual innovators to a broader speech community (Labov 2001). This step depends heavily on the innovators' relations to other speakers and the dynamics of group identity (Wedel and Van Volkinburg 2009).

In this account, phonological final devoicing is the end result of a diachronic process of phonologization (Hyman 1976), building on phonetic utterance-final devoicing. It is a hypocorrective sound change (Ohala 1981, 1993), beginning with a listener's failure to compensate perceptually for an effect of context on production.

Such a diachronic account would explain a number of the properties of phonological final devoicing. The phonological pattern is common and has emerged independently in numerous language groups because it results from a straightforward change based on a pervasive pattern of laryngeal coarticulation. Voiced obstruents are subject to change in final position because anticipation of the open glottis of nonspeech breathing diminishes the voicing that contributes to distinguishing those sounds. Voiced obstruents change to voiceless obstruents, because that is what a partially devoiced obstruent tends to be mistaken for. The pattern is restricted to obstruents, because partially devoiced sonorants are so low in intensity that they are mistaken for silence rather than for a voiceless segment (Myers and Hansen 2007). The sound change is recapitulated in first- and second-language acquisition because the identification errors that are the basis for the sound change are more frequent in inexperienced learners of the sound system.

The basic phonetic prerequisites for the phonological final devoicing pattern are widespread, perhaps universal. But not every language that has the phonetic pattern ends up with the corresponding phonological pattern. This is because the phonetic pattern of utterance-final devoicing is only the first step of the change, and the emergence of phonological final devoicing depends on all the subsequent steps as well. A listener has to make an identification error due to the phonetic pattern often enough that it serves as the basis of a generalization about voicing categories, and then this innovated

phonological pattern has to spread beyond that individual to a speech community. A sound change will only occur when these events happen to line up in the right way, but the point is that this series of events is more likely than one with a less frequent starting point (Yu 2004).

The articulatory and acoustic bases of this diachronic account are well-supported; the studies cited above demonstrate the effect of utterance-final position on the actions of the vocal folds and on the resulting acoustic reflexes of voicing. But no evidence has been provided to date for the claim that these acoustic effects of utterance-final coarticulation affect identification of voicing categories and lead to a tendency to identify utterance-final obstruents as voiceless. The aim of the present study is to test this claim.

English was chosen as the language of the study, since it has a robust contrast in voicing in word-final position (e.g. *pat/pad*), and English speakers are therefore experienced in producing and perceiving contrasting voice categories in this position. The first experiment is a production study, which is meant both to explore the conditions for the utterance-final devoicing effect and to generate stimuli for the perception experiments. The second and third experiments are perception experiments, in which English-speaking listeners identify words belonging to minimal pairs differing in final voicing (e.g. *proof/prove*), excised from utterance-final or nonfinal position.

## **1. Experiment #1: Production**

### **1.1 Methods**

The test items, listed in (2), all belonged to minimal pairs differing just in the voicing of a word-final obstruent: final voiceless vs. final voiced fricative (e.g. *proof/prove*), or final voiced vs. voiceless stops (e.g. *greet/greed*). There were ten test words for each class of word-final segment, for a total of 40 items.

(2) Test words

(a) Fricative-final (voiceless - voiced)

loose - lose  
proof - prove  
cease - seize  
leaf - leave  
Bruce - bruise  
noose - news  
belief - believe  
grief - grieve  
relief - relieve  
use (noun) - use (verb)

(b) Stop-final (voiceless - voiced)

greet - greed  
feet - feed  
beat - bead  
seat - seed  
loop - lube  
neat - need  
moot - mood  
leak - league  
heat - heed  
sweet - Swede

To control for the effects of stress (Lehiste 1970: 36) and vowel height (Lindblom 1968) on the duration of the vowel, the final syllable in all test words bore main word stress and had a high tense vowel.

Each test word occurred in two carrier sentences. To control for the effect of utterance length on vowel duration (Lindblom 1968), each carrier sentence consisted of 15 syllables. In one sentence, the test word was sentence-final (e.g. *The garage can tighten any of the bolts that are too **loose***), while in the other it was non-final and preceding a word beginning with a nasal stop (e.g. *There is a **loose** nylon cover over the whole area*). Each of the 40 test words occurred in 2 sentence positions, so there were 80 test sentences for each speaker.

6 adult native speakers of American English produced the materials.<sup>4</sup> The test sentences were presented in random order at 5-second intervals in a timed Powerpoint presentation on a laptop, interspersed with 60 distractor sentences (stimuli for another experiment). The speaker read them aloud seated in a sound recording booth and was recorded on a solid-state digital recorder. 480 test sentences were produced by the 6

---

<sup>4</sup> The subjects were from Oklahoma, Illinois, Tennessee, Oregon, and California.

speakers, but of those 1 item was excluded because the speaker produced the wrong word, and 2 were excluded due to background noise that made the measurements impossible. That left 477 items for analysis.

There are many acoustic correlates for voicing (Lisker 1986), but since this study concerns the interaction of such correlates with utterance position, we focus on measures which are defined for both stops and fricatives, and which are known to be sensitive to both voicing and utterance position.

The duration of the following intervals was measured, using Praat (<http://www.fon.hum.uva.nl/praat/>): (a) the vowel, (b) the constriction for the postvocalic consonant (if any), (c) the release (if any), and (d) the voiced and voiceless subintervals within the VC sequence (a) -(c).

The onset of the vowel was defined as the onset of an increase in amplitude and wave complexity after the prevocalic consonant. The offset of the vowel was defined as the end of F2 and F3 of the vowel. The constriction interval was a period of minimum amplitude beginning with the vowel offset. In the case of stops this corresponded to the closure interval, while in fricatives it was the noise interval. The release was the interval of increased amplitude after the constriction and before the onset of a following sound, if any. In a released stop, this was the noise burst and aspiration. In fricatives, this was a period characterized by a shift to lower frequency and lower intensity noise. In these recordings, all utterance-final stops were released, suggesting a fairly careful pronunciation.

The voiced interval was from the onset of the vowel to the offset of quasiperiodic pulses in the waveform. The voiceless interval consisted either of noise without such pulsing, or silence (in a stop closure). In these recordings, the voiced interval was always continuous, i.e. all the VC sequences measured had an initial voiced interval, which was followed in some cases by a voiceless interval that stretched to the end of the VC sequence. The transition from voiced to voiceless always occurred either in the constriction interval or in the release.

The constriction interval was expected to be longer in voiceless obstruents than in voiced obstruents (Denes 1955, Lisker 1957, Stevens et al. 1992), and longer in utterance-final than in nonfinal position (Byrd et al. 2005).

The duration of the voiceless interval within the consonant (constriction + release) was expected to be longer in a voiceless than in a voiced consonant (Raphael 1971, Kingston and Diehl 1994), and longer in utterance-final than in nonfinal position (Haggard 1978, Docherty 1992, Smith 1997). The absolute duration of the voiceless interval was chosen in this study as a measure rather than the proportion of the voiced interval to the whole consonant interval because the latter measure had a bimodal distribution, complicating statistical analysis (cf. Kuzla, Cho and Ernestus 2007). The duration of the voiceless interval was chosen over that of the voiced interval since it is a devoicing effect that we are aiming to measure.

The duration of the vowel was expected to be longer before voiced than before voiceless consonants (Chen 1970), and longer in utterance-final position than in nonfinal position (Lindblom 1968, Oller 1973). The ratio of the vowel duration to the duration of the whole vowel + consonant sequence has been argued to be a more robust acoustic correlate of voicing than the absolute duration of the vowel alone (Kohler 1979, Barry 1979, Pind 1986). This V/VC ratio is greater when C is voiced than when it is voiceless, because the vowel before a voiced coda is longer and the coda itself is shorter. The ratio is also lower in utterance-final syllables than in nonfinal syllables (Barry 1979), because final lengthening is gradient and has more of an effect on the utterance-final coda than on the preceding vowel, proportionally (Turk 1999).

In the statistical analysis a mixed model was used in which speaker and test word were treated as random effects, and fixed effects were voicing (voiced/voiceless), manner (stop/fricative), and (utterance) position (final/nonfinal). The alpha level was .05. A given acoustic measure was considered to show utterance-final devoicing when a significant effect of utterance-final position coincided with (i.e. went in the same direction as) the effect of having a voiceless obstruent in coda position.

## 1.2 Results

### 1.2.1 Constriction duration

Fig. 1 presents the percentile distribution of constriction duration by manner, voicing and position:

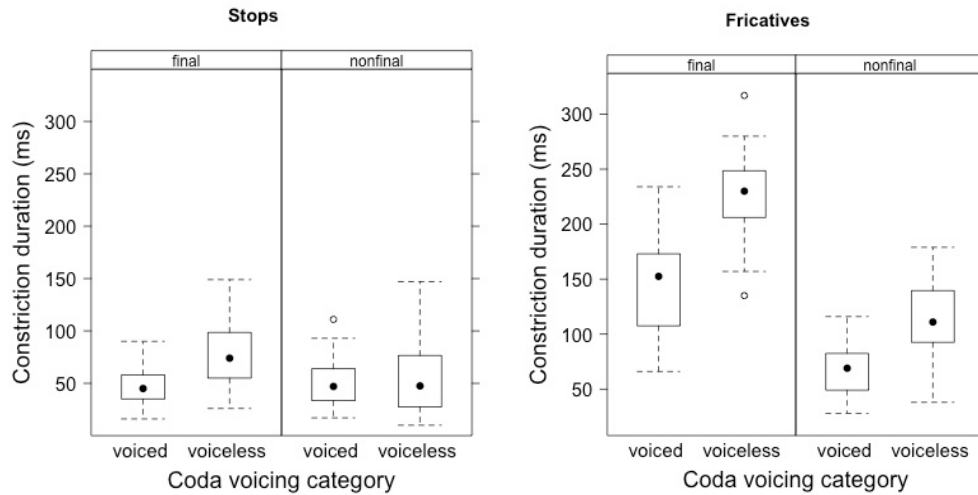


Fig. 1: Constriction duration (ms) by manner, voicing and position

Mean constriction duration was greater in voiceless obstruents (118 ms) than in voiced obstruents (77 ms), greater in fricatives (138 ms) than in stops (57 ms), and greater in utterance-final syllables (124 ms) than in nonfinal ones (71 ms). All three main effects were significant (d.f. = 1, 469): voicing ( $F = 109.0$ ,  $p < .001$ ), manner ( $F = 444.5$ ,  $p < .001$ ), and position ( $F = 433.4$ ,  $p < .001$ ). The 2-way interactions were also significant (d.f. = 1, 469): voicing\*position ( $F = 30.7$ ,  $p < .001$ ), voicing\*manner ( $F = 36.1$ ,  $p < .001$ ), and manner\*position ( $F = 274.7$ ,  $p < .001$ ). The 3-way interaction was not significant.

In order to explore the interactions, the data was split into subsets according to manner and position. Considering the manner classes, both the voicing and position main effects were significant in both stops and fricatives: voicing in stops ( $F(1, 234) = 17.0$ ,  $p < .001$ ), voicing in fricatives ( $F(1, 235) = 96.8$ ,  $p < .001$ ), position in stops ( $F(1, 234) = 13.1$ ,  $p < .001$ ), and position in fricatives ( $F(1, 235) = 612.1$ ,  $p < .001$ ). The interaction

between the two factors was significant in both stops ( $F(1, 234) = 17.5, p < .001$ ) and fricatives ( $F(1, 235) = 17.4, p < .001$ ).

Splitting the data according to position class, voicing was a significant main effect in both final and nonfinal positions (final,  $F(1, 233) = 190.9, p < .001$ ; nonfinal,  $F(1, 236) = 35.2, p < .001$ ). The main effect of manner was also significant in both position classes (final,  $F(1, 233) = 983.4, p < .001$ ; nonfinal,  $F(1, 236) = 78.7, p < .001$ ), as was the interaction of voicing and manner (final,  $F(1, 233) = 41.6, p < .001$ ; nonfinal,  $F(1, 236) = 23.1, p < .001$ ).

In both stops and fricatives, and in both final and nonfinal position, constriction duration was significantly longer in voiceless than in voiced obstruents, as previously found by Denes (1955). As in Byrd et al. (2005), the constriction duration was longer in utterance-final than in non-final position. The two factors interacted so that the difference between the voiced and voiceless group means was greater in final position (55 ms) than in nonfinal position (27 ms). The lengthening of constriction duration associated with utterance-final position coincided with and reinforced the lengthening effect of belonging to the voiceless category, so with respect to this measure utterance-final position had a devoicing effect. The same result is obtained if the whole consonant duration (constriction + release) is used as a measure, instead of just the constriction interval.

### **1.2.2 Duration of the voiceless interval**

Fig. 2 presents the percentile distribution of the duration of the voiceless interval by manner, voicing, and position class.

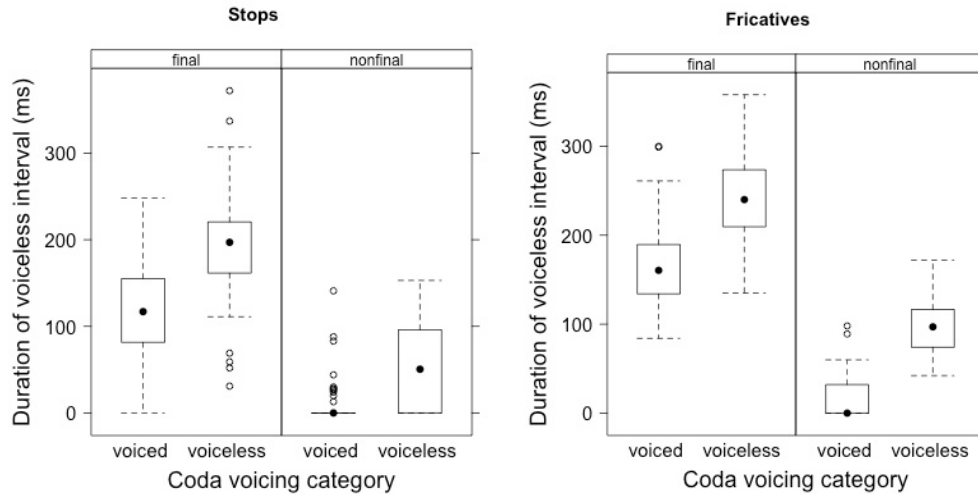


Fig. 3. Duration of the voiceless interval by manner, voicing and position

The mean duration of the voiceless interval was greater in voiceless obstruents (146 ms) than in voiced ones (78 ms), greater in fricatives (130 ms) than in stops (93 ms), and greater in final position (181 ms) than in nonfinal position (43 ms). All three main effects are significant (d.f. = 1, 469): voicing ( $F = 194.8$ ,  $p < .001$ ), manner ( $F = 56.9$ ,  $p < .001$ ) and position ( $F = 1213.4$ ,  $p < .001$ ). All of the interactions except that of voicing and position were significant (d.f. = 1, 469): voicing and manner ( $F = 5.0$ ,  $p = .03$ ), manner and position ( $F = 7.8$ ,  $p = .01$ ), and voicing\*manner\*position ( $F = 4.4$ ,  $p = .04$ ).

In the manner subsets, the main effect of voicing was significant (stops,  $F(1, 234) = 45.5$ ,  $p < .001$ ; fricatives,  $F(1, 235) = 272.4$ ,  $p < .001$ ), as was the main effect of position (stops,  $F(1, 234) = 406.8$ ,  $p < .001$ ; fricatives,  $F(1, 235) = 970.3$ ,  $p < .001$ ). The interaction of voicing and position was significant in stops ( $F(1, 234) = 5.0$ ,  $p = .03$ ), but not fricatives.

In the position subsets, the main effect of voicing was significant (final,  $F(1, 233) = 108.0$ ,  $p < .001$ ; nonfinal,  $F(1, 236) = 104.5$ ,  $p < .001$ ), and so was the main effect of manner (final,  $F(1, 233) = 44.8$ ,  $p < .001$ ; nonfinal,  $F(1, 236) = 17.4$ ,  $p < .001$ ). The interaction of voicing and manner was significant in the nonfinal subset ( $F(1, 236) = 9.7$ ,  $p = .002$ ), but not in the final subset.

Thus the voiceless interval in the consonant was significantly longer in voiceless obstruents than in voiced ones in both stops and fricatives, and both final and nonfinal

position. The voiceless interval was longer in utterance-final position than in nonfinal position. The voicing and position factors interacted in such a way that the difference between the voiced and voiceless group means was greater in final position (74 ms) than in nonfinal position (63 ms). For this measure, the lengthening effect of final position coincided with the lengthening effect of a voiceless obstruent, so utterance-final position had a devoicing effect. The same general pattern holds if the proportion of voicing to consonant duration is used as a measure.

### 1.2.3 Vowel duration

Fig. 3 presents the percentile distribution of vowel duration (ms) by coda voicing and utterance position:

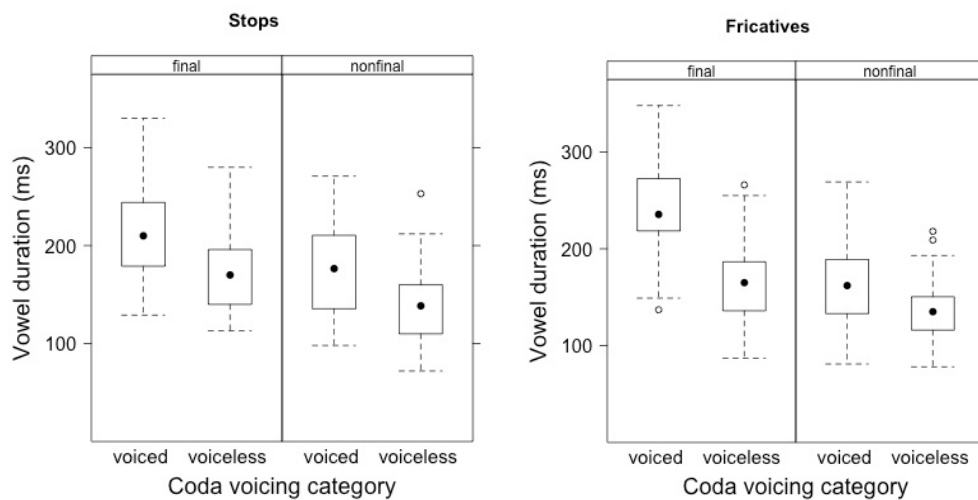


Fig. 3: Vowel duration (ms) by manner, voicing and utterance position

Mean vowel duration was greater before a voiced obstruent (199 ms) than before voiceless ones (mean = 152 ms), slightly greater before a fricative (177 ms) than before a stop (175 ms), and greater in an utterance-final syllable (198 ms) than in a nonfinal syllable (154 ms). The main effects of voicing and position were significant (d.f. = 1, 469): voicing,  $F = 57.8$ ,  $p < .001$ ; position,  $F = 255.0$ ,  $p < .001$ , but not the main effect of manner ( $F < 1$ ). The significant interactions (d.f. = 1, 469) were those of voicing and position ( $F = 21.0$ ,  $p < .001$ ), voicing and manner ( $F = 13.3$ ,  $p < .001$ ), and voicing,

manner and position ( $F = 18.8, p < .001$ ). The interaction of manner with voicing was not significant ( $F < 1$ ).

The main effect of voicing was significant in both manner subsets (stops,  $F(1, 234) = 22.7, p < .001$ ; fricatives,  $F(1, 235) = 35.6, p < .001$ ), as was the main effect of position (stops,  $F(1, 234) = 65.5, p < .001$ ; fricatives,  $F(1, 235) = 226.3, p < .001$ ). The interaction between voicing and position was significant in fricatives ( $F(1, 235) = 46.7, p < .001$ ), but not in stops ( $F < 1$ ).

Examining the final and nonfinal subsets separately, the main effect of voicing was significant in both the final and nonfinal subsets (final,  $F(1, 233) = 79.7, p < .001$ ; nonfinal,  $F(1, 236) = 19.2, p < .001$ ), while the main effect of manner was not significant in either subset. The interaction of voicing and manner was significant in the final subset (manner,  $F(1, 233) = 6.9, p = .009$ ), but not in the nonfinal subset.

Thus, as in previous studies (e.g. Chen 1970), vowel duration in this sample was significantly greater before a voiced consonant than before a voiceless one in stops and fricatives, and in final and nonfinal position. As in previous studies (e.g. Oller 1973), vowel duration was significantly greater in an utterance-final syllable than in a nonfinal one. The two factors interacted in that the difference between the voiced and voiceless group means was greater in final position (60 ms) than in nonfinal position (33 ms). Umeda (1975) found a similar interaction in a study of connected speech in English, but in her study the significant effects of voicing were limited to final position. For our purposes, what matters most is that with regard to this voicing cue, the lengthening effects of utterance-final position coincide with and reinforce the effect of a voiced coda, rather than the voiceless category as with the previous two measures. Thus with respect to this measure, utterance-final position has a voicing effect, not the expected devoicing effect.

However, it should be noted that some scholars have argued that the appropriate cue for voicing in a postvocalic consonant is not the absolute duration of the preceding vowel, but the ratio of the vowel duration to the duration of the whole VC sequence (Kohler 1979, Barry 1979, Pind 1986). The percentile distribution of this measure,  $V/VC$ , is presented in Fig. 4.

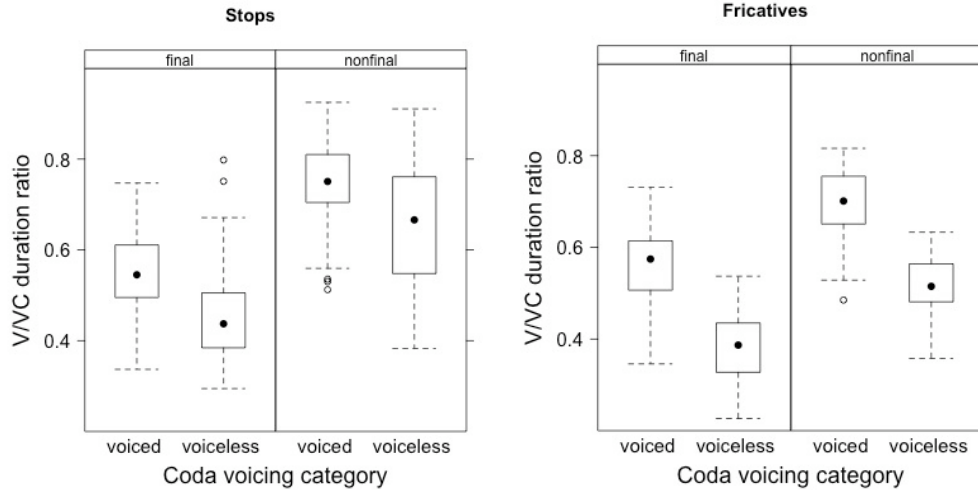


Fig. 4: The ratio of vowel duration to VC duration, by manner, voicing and position class

The mean V/VC ratio was greater for voiced codas (.64) than for voiceless ones (.51), greater with stops (.61) than with fricatives (.54), and greater in nonfinal position (.66) than in utterance-final position (.49). All the main effects were significant for this measure (d.f. = 1, 469): voicing ( $F = 129.2, p < .001$ ), manner ( $F = 36.9, p < .001$ ), and position ( $F = 568.8, p < .001$ ). The significant interactions (d.f. = 1, 469) were those of voicing and manner ( $F = 14.6, p < .001$ ), and manner and position ( $F = 25.8, p < .001$ ).

The main effect of voicing was significant in both the stop subset ( $F(1, 234) = 19.6, p < .001$ ) and the fricative subset ( $F(1, 235) = 220.3, p < .001$ ). The main effect of position was also significant in both manner subsets (stops,  $F(1, 234) = 325.9, p < .001$ ; fricatives,  $F(1, 235) = 269.0, p < .001$ ). In neither manner subset was there a significant interaction between voicing and position.

Dividing the dataset by position, the main effect of voicing was significant (final,  $F(1, 233) = 133.6, p < .001$ ; nonfinal,  $F(1, 236) = 60.4, p < .001$ ), as was that of manner (final,  $F(1, 233) = 8.5, p = .004$ ; nonfinal,  $F(1, 236) = 38.6, p < .001$ ) and the interaction of voicing and manner (final,  $F(1, 233) = 14.2, p < .001$ ; nonfinal,  $F(1, 236) = 7.1, p = .008$ ).

Thus the V/VC ratio was significantly greater with voiced codas than with voiceless ones in both stops and fricatives, and both final and nonfinal position. The ratio was significantly greater in final than in nonfinal position for both stops and fricatives. In

this case there was no significant interaction between these two factors, so that there was no significant difference in the voicing effect depending on utterance position. The lowered ratio associated with final position coincided with the lowering effect of a voiceless coda - an utterance-final devoicing effect.

#### **1.2.4 Discussion: Production study**

It has been found that there are three important acoustic cues for voicing that show utterance-final devoicing in this dataset, such that the effect of final position coincides with the effect of voicelessness. Constriction duration and the duration of the voiceless interval are longer in final than nonfinal position, and longer in voiceless than in voiced consonants. The V/VC duration ratio is lower in final than in nonfinal position, and lower with voiceless than with voiced codas.

Vowel duration goes the other direction, since vowels were longer in final position and with voiced codas. There are also other cues for voicing that have not been measured in this study (Lisker 1986), such as  $f_0$  or F1, in which the effects of final position might coincide with those of the voiced category.

The question is which class of acoustic correlates is dominant in the identification of voicing categories in utterance-final position. If, as we have hypothesized, listeners tend to identify utterance-final obstruents as voiceless, this would suggest that the cues that are most salient to listeners are those in which the effect of utterance-final position coincides with the effect of a voiceless coda. If, on the other hand, absolute vowel duration is the most salient acoustic correlate of coda voicing for listeners, then we would expect that listeners would tend to identify utterance-final obstruents as voiced. The perception experiments presented in the next session will provide evidence on this point.

## **2. Experiments #2a and #2b: Perception**

We have seen in the production study that utterance-final position has significant effects on the acoustic correlates of the voicing contrast in English. The goal of the

perception studies is to investigate what consequences, if any, these acoustic effects have on listeners' identification of categories contrasting in voicing.

## 2.1 Methods

The 20 obstruent-final voiced-voiceless pairs in (2) above served as the test items for the perception experiments. The test words were excised from the soundfiles produced in the first experiment, cutting at the closest zero crossing preceding the onset of the word and the closest zero crossing immediately following the end of the word (including the final consonant release, if any).

There were 477 tokens from the production study. 2 of these were excluded due to speaker error, leaving 475 stimuli. The recordings were processed in Adobe Audition. They were normalized to the same peak intensity (-15 dB relative to full scale). To decrease the abruptness of the soundfile onset, a 100 ms. interval of silence was added to the beginning of each soundfile, and if the initial sound wasn't a plosive, the initial portion was reset to fade in gradually.

In the stimuli for Experiment 2a, the files were presented in the clear, without any added noise. However, it was expected that the error rate would be low in this case, and perhaps too low to provide enough information about the kinds of errors listeners were prone to. Thus a new series of stimuli was created by taking the stimuli for Experiment 2a and mixing in pink noise (in which intensity is inversely proportional to frequency) at a signal-to-noise ratio of 10 dB. These stimuli were used in Experiment 2b.

Since no changes were made to  $f_0$ , voice quality, or segment duration, the stimuli do not sound like isolation words, i.e. complete intonational phrases consisting of a single word, but are clearly incomplete snippets from a longer utterance. Subjects were told that the stimuli were cut out of longer sentences, and that that is why they might sound odd.

The stimuli were presented through headphones from a laptop computer using Superlab (Cedrus). The stimuli were blocked by word pair, and presented in a different random order within the block for each subject. For each block, the subject was presented with the choice of items on the laptop screen – one choice in blue on the left side of the screen, and the other in red on the right. Sound files were presented every 2 seconds, and

subjects were instructed to listen to each one and press a key on a response box to indicate as quickly as possible their choice as to which word they heard. A blue key on the left of the box corresponded to the left-hand choice in blue, and a red key on the right corresponded to the right-hand choice in red. Voiced and voiceless choices were evenly distributed between left and right.

16 adult native speakers of American English participated in each of the two experiments. Each subject participated in only one of the experiments, and none of the subjects for the production experiment participated as subjects in these perception experiments.

Since the experiment involves native speakers of English identifying familiar words of English, it was expected that the error rate would be quite low. But the hypothesis was that there would nevertheless be a significant tendency among listeners to identify utterance-final obstruents as voiceless, and that there would be more voiceless judgements for forms excised from utterance-final position than for forms from utterance-medial position.

The dependent variable here is a categorical response (voiced/voiceless), so the models are based on a binomial distribution. Subject (i.e. listener), talker, and minimum pair were included as random effects, and the fixed effects were stimulus voicing (voiced coded as 0, and voiceless as 1) and stimulus position (final coded as 0, and nonfinal as 1). The logistic regression models expressed the likelihood of a voiceless response based on these factors.

## **2.2 Results**

There were 7600 trials in each experiment (475 stimuli \* 16 subjects). In experiment 2a there were 67 nonresponses, i.e. cases in which the subject did not press either key in the 2-second interval allowed. These were excluded from the analysis, leaving a total of 7533 responses. In Experiment 2b, 62 nonresponses were excluded, as well as 2 responses with suspiciously low response times below 110 ms (from stimulus onset), leaving a total of 7536 responses.

As expected, subjects were in general quite accurate in their identification of the stimuli. 6844 of the responses in Experiment 2a (90.9%) were correct, while in Experiment 2b, with added noise, 6539 responses (86.8%) were correct.

The responses are broken down by stimulus category in Table 1 for Experiment 2a and in Table 2 for Experiment 2b. Correct responses are highlighted in boldface.

**Table 1: Experiment 2a: Number (and percentage) of voicing responses by stimulus manner, position and voicing**

<b>Stimulus manner</b>	<b>Stimulus position</b>	<b>Stimulus voicing</b>	<b>Number (and percentage) of voiced responses</b>	<b>Number (and percentage) of voiceless responses</b>
<i>Stop</i>	<i>Nonfinal</i>	<i>Voiced</i>	<b>866 (92%)</b>	74 (8%)
<i>Stop</i>	<i>Nonfinal</i>	<i>Voiceless</i>	98 (10%)	<b>841 (90%)</b>
<i>Stop</i>	<i>Final</i>	<i>Voiced</i>	<b>891 (95%)</b>	49 (5%)
<i>Stop</i>	<i>Final</i>	<i>Voiceless</i>	72 (8%)	<b>866 (92%)</b>
<i>Fricative</i>	<i>Nonfinal</i>	<i>Voiced</i>	<b>919 (97%)</b>	33 (3%)
<i>Fricative</i>	<i>Nonfinal</i>	<i>Voiceless</i>	189 (20%)	<b>760 (80%)</b>
<i>Fricative</i>	<i>Final</i>	<i>Voiced</i>	<b>820 (89%)</b>	106 (11%)
<i>Fricative</i>	<i>Final</i>	<i>Voiceless</i>	68 (7%)	<b>881 (93%)</b>

**Table 2: Experiment 2b: Number (and percentage) of voicing responses by stimulus manner, position and voicing**

<b>Stimulus manner</b>	<b>Stimulus position</b>	<b>Stimulus voicing</b>	<b>Number (and percentage) of voiced responses</b>	<b>Number (and percentage) of voiceless responses</b>
<i>Stop</i>	<i>Nonfinal</i>	<i>Voiced</i>	<b>802 (86%)</b>	136 (14%)
<i>Stop</i>	<i>Nonfinal</i>	<i>Voiceless</i>	114 (12%)	<b>825 (88%)</b>
<i>Stop</i>	<i>Final</i>	<i>Voiced</i>	<b>833 (88%)</b>	110 (12%)
<i>Stop</i>	<i>Final</i>	<i>Voiceless</i>	83 (9%)	<b>855 (91%)</b>
<i>Fricative</i>	<i>Nonfinal</i>	<i>Voiced</i>	<b>859 (90%)</b>	93 (10%)
<i>Fricative</i>	<i>Nonfinal</i>	<i>Voiceless</i>	184 (19%)	<b>766 (81%)</b>
<i>Fricative</i>	<i>Final</i>	<i>Voiced</i>	<b>767 (83%)</b>	162 (17%)
<i>Fricative</i>	<i>Final</i>	<i>Voiceless</i>	115 (12%)	<b>832 (88%)</b>

In both experiments, the fricatives show a higher percentage of voiceless responses for corresponding stimuli from final position than for those from nonfinal position. Thus in Experiment 2a, voiceless fricative stimuli were correctly identified as voiceless in 93% of the final cases but only 80% of the nonfinal ones. Voiced fricative stimuli in the same experiment were incorrectly identified as voiceless in 11% of the final cases, as compared to 6% of the nonfinal ones. The same generalization held for the fricative stimuli in Experiment 2b.

However, the stops did not show such a pattern. In both experiments, as with fricative stimuli, voiceless stops were correctly identified as voiceless more often in final position than in nonfinal position. For example, among stimuli with voiceless stops in Experiment 2a, 92% were identified as voiceless in final position, compared to 90% in nonfinal position. But, unlike the fricative stimuli, the incorrect identification of voiced stops as voiceless was slightly *less* frequent in final position than in nonfinal position. For example, in Experiment 2a, 8% of the final voiced stops were identified as voiceless, compared with 10% of the nonfinal voiced stops.

The results of the statistical analysis are given in Table 3 for Experiment 2a, and in Table 4 for Experiment 2b.

**Table 3. Experiment 2a: Fixed effects**

<i>Fixed effect</i>	<i>Estimated coefficient</i>	<i>z</i>	<i>p</i>
Voicing	5.2	29.0	<.001
Position	-1.4	-6.4	<.001
Manner	-0.9	-3.3	<.001
Voicing * Position	0.1	0.4	.66
Voicing * Manner	0.9	6.4	<.001
Position* Manner	1.8	6.4	<.001
Voicing*Manner*Position	-1.0	-2.6	.01

**Table 4. Experiment 2b: Fixed effects**

<i>Fixed effect</i>	<i>Estimated coefficient</i>	<i>z</i>	<i>p</i>
Voicing	3.9	27.7	<.001
Position	-0.7	-4.9	<.001
Manner	-0.5	-2.2	.03
Voicing * Position	0.1	0.6	.53
Voicing * Manner	0.9	4.2	<.001
Position* Manner	1.0	4.9	<.001
Voicing*Manner*Position	-0.8	-2.7	.007

In both experiments, all the main effects were significant. The positive coefficient for voicing, in conjunction with the significant z value, indicates that a voiceless stimulus (coded 1) had a significantly greater likelihood to be identified as voiceless than a voiced stimulus (coded 0) had. This just reflects the high accuracy of identification of voicing categories. The negative coefficient for position, in conjunction with the significant effect for that factor, means that a nonfinal stimulus (coded 1) had a significantly smaller

likelihood to be identified as voiceless than a final stimulus (coded 0). The negative coefficient for manner meant that a stop (coded 1) was less likely than a fricative (coded 0) to be identified as voiceless. However, there were significant interactions among these factors, complicating the interpretation.

The dataset was split into stop and fricative subsets so that the effects of voicing and position could be examined in these different manner subsets. The results of these tests are presented in tables 5-8.

**Table 5. Experiment 2a: Fricatives**

<i>Fixed effect</i>	<i>Estimated coefficient</i>	<i>z</i>	<i>p</i>
Voicing	5.3	28.1	< .001
Position	-1.4	-6.5	< .001
Voicing * Position	0.2	0.6	.53

**Table 6. Experiment 2b: Fricatives**

<i>Fixed effect</i>	<i>Estimated coefficient</i>	<i>z</i>	<i>p</i>
Voicing	4.0	27.1	< .001
Position	-0.7	-4.9	< .001
Voicing * Position	0.1	0.6	.53

**Table 7. Experiment 2a: Stops**

<i>Fixed effect</i>	<i>Estimated coefficient</i>	<i>z</i>	<i>p</i>
Voicing	6.1	27.8	< .001
Position	0.5	2.4	.02
Voicing * Position	-0.8	-3.2	.001

**Table 8. Experiment 2b: Stops**

<i>Fixed effect</i>	<i>Estimated coefficient</i>	<i>z</i>	<i>p</i>
Voicing	4.8	28.6	< .001
Position	0.3	2.0	.048
Voicing * Position	-0.7	-3.1	.002

In the fricative subset in both experiments (Tables 5 and 6), the main effect of voicing was significant, with a positive coefficient reflecting the subjects' largely accurate identification of voicing categories. The main effect of position was also significant, with a negative coefficient reflecting an association of final position with voiceless responses and nonfinal position with voiced responses. The interaction was not significant, so there was no evidence that the effect of final position on identification was different for stimuli with a voiced fricative compared to those with a voiceless fricative.

In the stop subset (Tables 7 and 8), both main effects and their interaction were significant. The coefficient for voicing was positive in both experiments, indicating that among stops, as with fricatives, voiceless stimuli were more likely to get voiceless responses than were voiced stimuli. But the coefficient for position is positive, and that for the interaction is negative. This reflects the fact, noted above, that the effect of position was different for voiced stimuli than for voiceless ones. Voiceless stops were more often correctly identified as voiceless in final than in nonfinal position, as expected and as found with fricatives, but voiced stops were unexpectedly more often identified as voiceless in nonfinal than in final position.

The nonfinal condition was expected to be a neutral control condition, and in Experiment 2b this is how it turned out, with voiced and voiceless responses evenly split. But in Experiment 2a there were more voiced responses than voiceless responses in nonfinal position (51% among the stops, and 58% among the fricatives). This tendency toward voiced responses could be due to coarticulatory voicing, since each test obstruent is between two voiced sonorants (a vowel and a nasal). But it raises the possibility that the observed effects of position are due more to nonfinal voicing than to final devoicing. To exclude this possibility, we restrict our view to utterance-final fricatives. In this set, in

both experiments, the proportion of voiceless responses was significantly greater than that expected by chance: Experiment 2a,  $z = 1.7$ ,  $p = .04$ ; Experiment 2b,  $z = 2.2$ ,  $p = .02$ .

### **2.3 Discussion: Perception Studies**

The hypothesis was that the acoustic effects of utterance-final devoicing observed in the production study would lead to a tendency for word-final obstruents from utterance-final position to be identified as voiceless. The two perception experiments have provided support for this hypothesis in fricative-final words, but not in stop-final words, and in particular not in words ending in voiced stops.

One explanation of this difference between stops and fricatives could lie in the effects of manner on the acoustic correlates of voicing found in the production study. Fricatives had a longer constriction duration, a longer voiceless interval, and a lower V/VC duration ratio than stops, and in all of these measures the effect of a fricative thus coincided with the effect of a voiceless sound.<sup>5</sup> As Ohala (1983: 201) pointed out, "voiced fricatives have more exacting aerodynamic requirements than do voiced stops", since voicing requires supralaryngeal pressure to be lower than sublaryngeal pressure, while at the same time supralaryngeal pressure behind the oral constriction must be high enough to generate turbulent oral airflow. Ohala suggests that this aerodynamic issue is part of the reason that consonant inventories with only voiceless fricatives are twice as common in Ruhlen's (1987) survey than inventories in which all stops are voiceless. He goes on to propose that it is also the reason that "in American English the 'voiced' fricatives /v, z/ are more likely to be devoiced in word-final position than are the stops /b, d, g/." Thus it could be that the inherently weaker voicing cues for fricatives are more susceptible to confusing interference from utterance-final devoicing than the stronger voicing cues for stops.

---

<sup>5</sup> Manner had no significant effect on the last measure considered: absolute vowel duration.

### 3. Conclusion

The production study reported here has replicated previous findings of a significant utterance-final devoicing effect in English. Utterance-final obstruents displayed a longer constriction duration, a longer voiceless interval, and a lower ratio of vowel duration to total VC duration in comparison to nonfinal obstruents. In all these measures, the effect of utterance-final position coincides with the effect of belonging to the voiceless category.

The two perception studies tested the hypothesis that the acoustic effects of utterance-final position lead to a tendency to identify utterance-final obstruents as voiceless. The results of the studies clearly support that hypothesis for fricatives, but not for stops. Voiceless stops in final position were more likely to be correctly identified as voiceless than those in nonfinal position, but there was no tendency in either perception study for voiced stops to be misidentified as voiceless. I have suggested that this difference might have been due to that the effects of fricatives on the voicing cues coincides with the effects of voicelessness and utterance-final position.

The results provide a basis for rejecting the null hypothesis that utterance-final position has no effect on identification of voicing categories. The fact that fricatives from utterance-final position tended to be identified as voiceless further suggests that for word-final fricatives the dominant perceptual cues for voicing are those like constriction duration, voiceless interval duration, and V/VC ratio, for which the effects of final position coincide with those of the voiceless category.

The test words in the perception experiments were presented in isolation, without following context. Thus listeners were not able to use their abilities to compensate for the acoustic effects of coarticulation (Lindblom and Studdert-Kennedy 1967, Mann and Repp 1980), and to fill in missing information top-down from discourse context (Warren 1970) or from knowledge of the lexicon (Ganong 1980). These abilities are impressive, but they are not infallible, as evidenced by the fact that listeners do make identification errors in conversation even with the full phonetic and discourse context known. The

perception experiments reported here provide information about a baseline pattern of errors when such contextual information is held constant by complete removal of context.

The results thus provide support, at least in the case of fricatives, for one step in the diachronic account sketched in the introduction of how utterance-final phonetic devoicing provides the basis for a sound change resulting in phonological final devoicing. In this account the phonetic basis of final devoicing is limited to utterance-final position, but the pattern is generalized from there to word-final and syllable-final position through analogical extension. The results of the perception experiments also suggest that the perceptual basis of the sound change might be limited to fricatives, and extended from that subset of obstruents to the class of all obstruents.

One might expect from this that utterance-final fricative devoicing should be the most common version of the phonological pattern of final devoicing, since it requires the fewest further steps of generalization. There are cases of devoicing limited to fricatives (e.g. Gothic: Wright 1899: 62-67; Hock 1991: 43) and there are cases in which the devoicing is said to be limited to utterance-final position (e.g. some Yiddish dialects: Wetzels and Mascaró 2001: 224; Polish: Jassem and Richter 1989; examples in Blevins 2006: 142). But it certainly does not seem as if such cases are more common than word-final devoicing, or devoicing of all obstruents including stops.

It would appear, then, that the tendency to generalize the phonological pattern, from utterance-final words to all words and from fricatives to all obstruents is strong enough to render utterance-final phonological fricative devoicing unstable. If so, this tendency must lie not in the phonetic basis of the pattern, but in how language learners make generalizations about the distribution of speech sound categories (Hayes 1999, Moreton 2008).

## References

- Abu-Mansour, M. (1996). Voice as a privative feature: Assimilation in Arabic. In M. Eid (ed.) *Perspectives on Arabic Linguistics VIII*. John Benjamins, Amsterdam. 201-231.
- Ahmad, Z. (2005). *The Phonology-Morphology Interface in Malay: An Optimality Theoretic Account*. Pacific Linguistics, Canberra.

- Bach, E. (1996). Building words in Haisla. Unpublished manuscript, University of Massachusetts, Amherst.
- Baroni, M. and L. Vanelli (2000). The relationship between vowel length and consonantal voicing in Friulian. In L. Repetti (ed.) *Phonological Theory and the Dialects of Italy*. John Benjamins, Amsterdam. 13-44.
- Barry, W. (1979). Complex encoding in word-final voiced and voiceless stops. *Phonetica* 36. 361-372.
- Blevins, J. (2006). A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics* 32. 117-166.
- Bliese, L. (1981). *A Generative Grammar of Afar*. Summer Institute of Linguistics, Arlington (Texas).
- Bloomfield, L. (1933). *Language*. Holt, Rinehart, and Winston, New York.
- Booij, G. (1995). *The Phonology of Dutch*. Oxford University Press, Oxford.
- Broselow, E., S.-I. Chen, and C. Wang (1998). The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition* 20. 261-280.
- Byrd, D., S. Lee, D. Riggs, and J. Adams (2005). Interacting effects of syllable and phrase position on consonant articulation. *Journal of the Acoustical Society of America* 118. 3860-3873.
- Charles-Luce, J. and D. Dinnsen (1987). A reanalysis of Catalan devoicing. *Journal of Phonetics* 15. 187-190.
- Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22, 129-159.
- Clements, G. N. and S. J. Keyser (1983). *CV Phonology*. MIT Press, Cambridge.
- Cook, T. L. (1969). *The Pronunciation of Efik for Speakers of English*. Intensive Language Training Center, Indiana University, Bloomington.
- Cutts, H. and P. Jensen (1983). Speech timing of phonologically disordered children: Voicing contrast of initial and final stop consonants. *Journal of Speech and Hearing Science* 26. 501-510.
- Dambriunas, L., A. Klimas, and W. Smalstie (1966). *Introduction to Modern Lithuanian*. Franciscan Fathers, Brooklyn.

- Denes, P. (1955). Effect of duration on the perception of voicing. *Journal of the Acoustical Society of America* 27. 761-764.
- Dinnsen, D. and J. Charles-Luce (1984). Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics* 12. 49-60.
- Dinnsen, D. and F. Eckman (1978). Some substantive universals in atomic phonology. *Lingua* 45. 1-14.
- Dmitrieva, O. (2005). *Incomplete Neutralization in Russian Final Devoicing: Acoustic Evidence from Native Speakers and Second Language Learners*. Master's thesis, University of Kansas, Lawrence.
- Docherty, G. (1992). *The Timing of Voicing in British English Obstruents*. Foris, Berlin.
- Eckman, F. (1981). On the naturalness of interlanguage phonological rules. *Language Learning* 31. 195-216.
- Ernestus, M. and R. H. Baayen (2006). The functionality of incomplete neutralization in Dutch: The case of past-tense formation. In L. Goldstein, D. Whalen, and C.T. Best (eds.) *Laboratory Phonology* 8. Berlin, Mouton de Gruyter. 27-49.
- Ewert, A. (1933). *The French Language*. Faber and Faber, London.
- Flege, J. E. (1982). English speakers learn to suppress final devoicing. In K. Tuite, R. Schneider, and R. Chametsky (eds.) *Papers from the 18<sup>th</sup> Regional Meeting*. Chicago Linguistics Society, Chicago. 111-122.
- Flege, J. and R. Davidian (1984). Transfer and developmental processes in adult foreign language speech production. *Applied Psycholinguistics* 5. 323-347.
- Flege, J., M. Munro, and L. Skelton (1992). Production of the word-final English /t/ - /d/ contrast by native speakers of English, Mandarin, and Spanish. *Journal of the Acoustical Society of America* 92. 128-143.
- Francard, M. and Y.-C. Morin (1986). Sandhi in Walloon. In H. Andersen (ed.) *Sandhi Phenomena in the Languages of Europe*. Mouton de Gruyter, Berlin. 453-474.
- Ganong, W. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance* 6. 110-125.
- Grammont, M. (1933). *Traité de phonétique*. Librairie Delagrave, Paris.
- Haggard, M. (1978). The devoicing of voiced fricatives. *Journal of Phonetics* 6. 95-102.
- Hahn, R. (1991). *Spoken Uyghur*. University of Washington Press, Seattle.

- Harris, J. (2009). Why final devoicing is weakening. In K. Nasukawa & P. Backley (eds.), *Strength Relations in Phonology*. Mouton de Gruyter, Berlin. 9-46.
- Hayes, B. (1999). Phonetically-driven phonology: The role of Optimality Theory and inductive grounding. In M. Darnell, E. Moravcsik, M. Noonan, F. Newmeyer, and K. Wheatly (eds.) *Functionalism and Formalism in Linguistics, Volume I: General Papers*. John Benjamins, Amsterdam. 243-285.
- Heim, M. (1976). *Contemporary Czech*. Slavica Publishers, Columbus, Ohio.
- Hock, H. H. (1991). *Principles of Historical Linguistics*. 2<sup>nd</sup> ed. Mouton de Gruyter, Berlin.
- Hodson, B. and E. P. Paden (1981). Phonological processes which characterize unintelligible and intelligible speech in early childhood. *Journal of Speech and Hearing Disorders* 46. 369-373.
- Hogan, J. and A. Roszypal (1980). Evaluation of vowel duration as a cue for the voicing distinction in the following word-final consonant. *Journal of the Acoustical Society of America* 67. 1764-1771.
- Holmes, J. (1996). Losing voice: Is final devoicing a feature of Maori English? *World Englishes* 15. 193-205.
- Hualde, J. (1991). *Basque Phonology*. Routledge, London.
- Hualde, J. (1992). *Catalan*. Routledge, London.
- Hyman, L. (1976). Phonologization. In A. Juillard (ed.) *Linguistic Studies Offered to Joseph Greenberg, v. 2*. Anma Libri, Saratoga, Ca., 407-418.
- Ingram, D. (1989). *Phonological Disability in Children*. Whurr, London.
- Iwasaki, S. and B. P. Ingkaphirom (2005). *Thai*. Cambridge University Press, Cambridge.
- Jakobi, A. (1990). *A Fur Grammar*. Helmut Buske, Hamburg.
- Jassem, W., and L. Richter (1989). Neutralization of voicing in Polish obstruents. *Journal of Phonetics* 17. 317-325.
- Jespersen, O. (1926). *Lehrbuch der Phonetik*. B. G. Truebner, Leipzig.
- Jessen, M. (1998). *Phonetics and Phonology of Tense and Lax Obstruents in German*. John Benjamins, Amsterdam.
- Jessen, M. and C. Ringen (2002). Laryngeal features in German. *Phonology* 19. 189-218.
- Jungraithmayr, H. (1970). *Die Ron-Sprachen*. Verlag J. J. Augustin, Glückstadt.

- Kingston, J. and R. Diehl (1994). Phonetic knowledge. *Language* 70. 419-454.
- Klatt, D. and L. Klatt (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America* 87. 820-857.
- Kohler, K. (1979). Dimensions in the perception of fortis and lenis plosives. *Phonetica* 36. 332-343.
- Krämer, M. (2000). Voicing alternations and underlying representations: The case of Breton. *Lingua* 110. 639-663.
- Kuzla, C., T. Cho, and M. Ernestus (2007). Prosodic strengthening of German fricatives in duration and assimilatory devoicing. *Journal of Phonetics* 35. 301-320.
- Labov, W. (2001). *Principles of Linguistic Change: Social Factors*. Blackwell, Malden (Mass.)
- Lehiste, I. (1970). *Suprasegmentals*. MIT Press, Cambridge.
- Lehtonen, J. (1970). *Aspects of Quantity in Standard Finnish*. Jyväskylä University Press, Jyväskylä, Finland.
- Lightner, T. (1972). *Problems in the Theory of Phonology. Vol. 1: Russian Phonology and Turkish Phonology*. Linguistic Research, Inc., Edmonton.
- Lindblom, B. (1968). Temporal organization of syllable production. *Speech Transmission Laboratory Quarterly Progress and Status Report 2-3* (Royal Institute of Technology, Stockholm). 1-6.
- Lindblom, B. (1983). Economy of speech gestures. In P. MacNeilage (ed.) *The Production of Speech*. Springer, New York. 217-245.
- Lindblom, B. and M. Studdert-Kennedy (1967). On the role of formant transitions in vowel recognition. *Journal of the Acoustical Society of America* 42. 830-843.
- Lisker, L. (1957). Closure duration and the intervocalic voiced-voiceless distinction in English. *Language* 33. 42-49.
- Lisker, L. (1986). 'Voicing' in English: A catalogue of acoustic features signaling /b/ vs. /p/ in trochees. *Language and Speech* 29. 3-11.
- Lisker, L., A. Abramson, F. Cooper, and M. Schvey. (1969). Transillumination of the larynx in running speech. *Journal of the Acoustical Society of America* 45. 1544-1546.

- Locke, J. (1983). *Phonological Acquisition and Change*. Academic Press, New York.
- Lombardi, L. (1995). Laryngeal neutralization and syllable well-formedness. *Natural Language and Linguistic Theory* 13. 39-74.
- Lombardi, L. (1999). Positional faithfulness and voicing assimilation in Optimality Theory. *Natural Language and Linguistic Theory* 17. 267-302.
- Luelsdorf, P. (1975). *A Segmental Phonology of Black English*. Mouton, The Hague.
- Major, R. and M. Faudree (1996). Markedness universals and the acquisition of voicing contrasts by Korean speakers of English. *Studies in Second Language Acquisition* 18. 69-90.
- Mann, V. and B. Repp (1980). Influence of vocalic context on perception of the [ʃ]-[s] distinction. *Perception and Psychophysics* 28, 213-228.
- Michelson, K. (1988). *A Comparative Study of Lake-Iroquoian Accent*. Kluwer, Dordrecht.
- Montreuil, J.-P. (1999). The Romansch syllable. In H. van der Hulst and N. Ritter (eds.) *The Syllable: Views and Facts*. Mouton de Gruyter, Berlin. 527-550.
- Moreton, E. (2008). Analytic bias and phonological typology. *Phonology* 25. 83-127.
- Myers, S. (2005). Vowel duration and neutralization of vowel length contrasts in Kinyarwanda. *Journal of Phonetics* 33. 427-446.
- Myers, S. and B. Hansen (2007). The origin of vowel length neutralization in final position: Evidence from Finnish speakers. *Natural Language and Linguistic Theory* 25.157-193.
- Newton, B. (1972). *The Generative Interpretation of Dialect: A Study of Modern Greek Phonology*. Cambridge University Press, Cambridge.
- Nivens, R. (1992). A lexical phonology of West Tarangan. In D. Burquest and W. Laidis (eds.) *Phonological Studies in Four Languages of Maluku*. Summer Institute of Linguistics, Arlington, Texas. 127-227.
- Ohala, J. (1981). The listener as a source of sound change. In C. Masek, R. Hendrick and M. Miller (ed.) *Papers from the Parasession on Language and Behavior*. Chicago Linguistics Society, Chicago. 178-203.
- Ohala, J. (1983). The origin of sound patterns in vocal tract constraints. In P. MacNeilage (ed.) *The Production of Speech*. Springer, New York. 189-216.

- Ohala, J. (1993). The phonetics of sound change. In C. Jones (ed.) *Historical Linguistics: Problems and Perspectives*. Longman, London, 237-278.
- Oller, D. (1973). The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America* 54. 1235-1247.
- Oller, D. and B. Smith (1977). Effect of final syllable position on vowel duration in infant babbling. *Journal of the Acoustical Society of America* 62. 994-997.
- Oller, D., L. Wieman, W. Doyle and C. Ross (1976). Infant babbling and speech. *Journal of Child Language* 3. 1-11.
- Padgett, J. (2002). Russian voicing assimilation, final devoicing, and the problem of [v]. Unpublished paper available at <http://people.ucsc.edu/~padgett/papers.html>.
- Passy, P. (1891). *Étude sur les changements phonétiques et leurs caractères généraux*. Librairie Firmin-Didot, Paris.
- Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee and P. Hopper (eds.) *Frequency and the Emergence of Linguistic Structure*. John Benjamins, Amsterdam. 137-157.
- Pind, J. (1986). The perception of quantity in Icelandic. *Phonetica* 43. 116-139.
- Pitkin, H. (1984). *Wintu Grammar*. University of California Press, Berkeley.
- Poppe, N. (1960). *Buriat Grammar*. Mouton, The Hague.
- Port, R. and M. O'Dell (1985). Neutralization of syllable-final devoicing in German. *Journal of Phonetics* 13, 455-471.
- Raphael, L. (1971). Preceding vowel duration as a cue to the perception of the voicing characteristics of word-final consonants in American English. *Journal of the Acoustical Society of America* 51. 1296-1303.
- Rubach, J. (1984). *Cyclic and Lexical Phonology: The Structure of Polish*. Foris, Dordrecht.
- Rubach, J. (1993). *The Lexical Phonology of Slovak*. Oxford University Press, Oxford.
- Ruhlen, M. (1987). *A Guide to the Languages of the World*. Stanford University Press, Stanford.
- Sapir, E. (1990). *Takelma Texts and Grammar (Collected Works of Edward Sapir VIII)*. Mouton de Gruyter, Berlin.

- Scatton, E. (1984). *A Reference Grammar of Modern Bulgarian*. Slavica Publishers, Columbus, Ohio.
- Selkirk, E. (1978). On prosodic structure and its relation to syntactic structure. In T. Fretheim (ed.) *Nordic Prosody II*. TAPIR, Trondheim. 111-140.
- Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology Yearbook* 3. 371-405.
- Shadle, C. (1997). The aerodynamics of speech. In W. Hardcastle and J. Laver (eds.) *The Handbook of Phonetic Sciences*. Blackwell, Oxford. 33-64.
- Sievers, E. (1901). *Grundzüge der Phonetik*. 5th ed. Breitkopf und Härtel, Leipzig.
- Singh, C. Y. (2000). *Manipuri Grammar*. Rajesh Publications, New Delhi.
- Slifka, J. (2006). Some physiological correlates to regular and irregular phonation at the end of an utterance. *Journal of Voice* 20. 171-186.
- Slowiaczek, L. and D. Dinnsen (1985). On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics* 13. 325-341.
- Slowiaczek, L. and H. Szymanska (1989). Perception of word-final devoicing in Polish. *Journal of Phonetics* 17. 205-212.
- Smith, B. (1979). A phonetic analysis of consonantal devoicing in children's speech. *Journal of Child Language* 6. 19-28.
- Smith, C. (1997). The devoicing of /z/ in American English: Effects of local and prosodic context. *Journal of Phonetics* 25. 471-500.
- Smith, C. (1999). Marking the boundary: Utterance-final prosody in French questions and statements. In J. Ohala, Y. Hasegawa, M. Ohala, D. Granville and A. Bailey (eds.), *Proceedings of the XIVth International Congress of Phonetic Sciences*. Linguistics Department, University of California: Berkeley. 1181-1184.
- Smith, C. (2003). Vowel devoicing in contemporary French. *French Language Studies* 13. 177-194.
- Smith, N. (1973). *The Acquisition of Phonology*. Cambridge University Press, Cambridge.
- Stevens, K., S. Blumstein, L. Glicksman, M. Burton, and K. Kurowski (1992). Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. *Journal of the Acoustical Society of America* 91. 2979-3000.

- Sweet, H. (1877). *Handbook of Phonetics*. Oxford University Press, Oxford.
- Thompson, L. (1965). *A Vietnamese Grammar*. University of Washington Press, Seattle.
- Thurgood, G. and R. LaPolla (eds.) (2003). *The Sino-Tibetan Languages*. Routledge, London.
- Tucker, A. (1994). *A Grammar of Kenya Luo (Dholuo)*. Rüdiger Koppe Verlag, Köln.
- Turk, A. (1999). Structural influences on boundary-related lengthening in English. In J. Ohala, Y. Hasegawa, M. Ohala, D. Granville and A. Bailey (eds.), *Proceedings of the XIVth International Congress of Phonetic Sciences*. Linguistics Department, University of California, Berkeley. 237-240.
- Umeda, N. (1975). Vowel duration in American English. *Journal of the Acoustical Society of America* 58. 434-445.
- Velten, H. (1943). The growth of phonemic and lexical patterns in infant speech. *Language* 19. 281-292.
- Warner, N., Jongman, A., Sereno, J. and Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics* 32. 251-276.
- Warren, R. (1970). Perceptual restoration of missing speech sounds. *Science* 167, 392-393.
- Wedel, A. and H. Van Volkinburg (2009). Modeling simultaneous convergence and divergence of linguistic features between differently-identifying groups in contact. Unpublished paper downloaded from <http://dingo.sbs.arizona.edu/~wedel/publications/>.
- Westbury, J. and P. Keating (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics* 22. 145-166.
- Wetzels, W. and J. Mascaró (2001). The typology of voicing and devoicing. *Language* 77. 207-244.
- Whitney, W. D. (1879) *A Sanskrit Grammar*. Breitkopf and Härtel, Leipzig.
- Wolf, C. (1978). Voicing cues in English final stops. *Journal of Phonetics* 6. 299-309.
- Wolfram, W. (1969). *A Sociolinguistic Description of Detroit Negro Speech*. Center for Applied Linguistics, Washington, D.C.

Wolfram, W., and D. Christian (1976). *Appalachian Speech*. Center for Applied Linguistics, Arlington, Va.

Wright, J. (1899). *A Primer of the Gothic Language*. Oxford University Press, Oxford.

Yu, A. (2004). Explaining final obstruent voicing in Lezgian: Phonetics and history. *Language* 80. 73-97.