

Keeping and Losing Contrasts¹

JOHN KINGSTON

University of Massachusetts, Amherst

0. Introduction

In this paper, I will first show that contrasts aren't licensed in particular contexts because they are necessarily realized with more salient cues in those contexts than in contexts where they aren't licensed. Next, I will argue that the releases of consonants, particularly non-continuants, are not packages of particularly salient acoustic information about the consonant's identity but instead only one of many kinds of acoustic information about how the utterance is segmented prosodically. Prosodic segmentation helps the listener find words in the stream of speech. Finally, I will show that segments are usually perceived to be different from their neighbors, i.e. to contrast with them, except when the target sound is C_1 in a VC_1C_2V string, which is instead often perceived to be the same as C_2 , i.e. to assimilate to C_2 . Both the general and specific effects of neighboring sounds phonetically explain where contrasts are kept or lost quite differently than the licensing by cue account.

1. Licensing by cue

Licensing by cue explains where phonological contrasts are kept or lost in terms of where they are or are not expressed by perceptually salient cues (Steriade 1995, 2000, 2001; Wilson 2001). The explanation rests on two incorrect assumptions. First, that contrasts are pronounced in the same way, and thus would be recognized by the same potential cues, in all contexts, but that these cues differ in salience or even audibility between contexts. And second, that the cues themselves differ in their inherent perceptual salience. The first assumption is incorrect because contrasts are typically pronounced differently in different contexts, and these differences may be designed to optimize conveying them in each context. The second assumption is incorrect because languages differ in

¹ I am grateful to the audiences at the 28th Meeting of the Berkeley Linguistic Society, the University of Massachusetts, Amherst, and the University of Pennsylvania for comments and questions that have substantially improved this paper. Maciej Baranowski, Gene Buckley, Ed Flemming, Mark Liberman, John McCarthy, Pawel Nowak, and Bob Rothstein deserve special thanks. All errors are mine.

which cues they select to convey the same contrast, and which cues they transform or transfer the contrast to in the course of sound change.

1.1. Salience differences between contexts and languages?

Kingston & Diehl (1994; also Keating (1984), Maddieson (this volume)) have shown that speakers pronounce the members of contrasts differently in different contexts, languages, and styles, and that as a result, the acoustic differences between these contrasting sounds aren't always the same. A familiar example of such variation across contexts is that of the [voice] contrast in English stops (Lisker 1986, Kingston & Diehl 1994):

(1) English stops contrasting for [voice] differ:

(a) Word-initially and pretonically in:

(i) Aspiration

(b) Intervocally and not pretonically in:

(i) Closure duration: /b,d,g/ < /p,t,k/

(ii) Preceding vowel duration: /b,d,g/ > /p,t,k/

(iii) Closure voicing: /b,d,g/ > /p,t,k/

(iv) F₁ at vowel edge: /b,d,g/ < /p,t,k/

(v) F₀ at vowel edge: /b,d,g/ < /p,t,k/

(c) Word- and syllable-finally in:

(i) Preceding vowel duration

(ii) Glottalization

(More detail is given for intervocalic stops because the perceptual value of these differences is discussed below.)

Laeufer (1992) documents particularly interesting differences and similarities across languages in the realization of the word-final [voice] contrast in English and French. Since Chen's 1970 survey, received wisdom had been that vowels lasted 50% or more longer before [+voice] than [-voice] stops in English but only 10-20% longer in French and other languages (also Mack (1982)). Laeuffer showed instead that the vowel duration differences were much more similar to one another in these two languages so long as the stops are syllabified in the same way, which is difficult to achieve given their very different prosodies. When prosody re-syllabifies the word-final consonant into the following syllable's onset in either language, preceding vowels differ much less in duration than when the prosody leaves the stop in the preceding syllable's coda. Moreover, in French, final [+voice] and [-voice] obstruents are very often released into voiced and voiceless vocoids, respectively, and these vocoids behave like genuine vowels in permitting the final consonants to re-syllabify. Because these distinctive releases are so frequent, there are many more re-syllabified tokens in French than in English, and consequently preceding vowel durations differ less over all tokens in

French than in English. In other words, French speakers rely more on the distinctive consonant releases to convey a final [voice] contrast than English speakers do, but when they don't pronounce the stops with distinctive releases, differences in preceding vowel duration appear that are comparable in size to those observed in English. (Davis & Summers (1989) also show that vowel duration differences depend on syllabification in English, being much larger post-tonically, in words like *rabid* vs. *rapid*, than pre-tonically, in words like *rebel* (verb) vs. *repel*.)

Kingston & Diehl also argued that in some cases speakers choose different pronunciations to optimize conveying the contrast in each context that it occurs. The argument has three parts. First, all contrasts are conveyed by many, covarying acoustic differences. Second, many of these acoustic differences are produced by independently controlled articulations, not by some articulations occurring mechanically as a consequence of another controlled articulation. (As Kingston & Diehl defend this claim extensively, it won't be discussed further here.) Third, covarying acoustic differences often integrate perceptually so as to enhance the contrast.

Perceptual integration has been studied most thoroughly for the [voice] contrast in intervocalic stops, where it produces three perceptual properties: the low (acoustic) frequency property, the closure duration property, and the consonant:vowel duration ratio.

Low Frequency. Kingston & Diehl (1995) and Kingston, *et al.* (submitted) found that two stops were easy to discriminate when one had low F_1 or F_0 at the edges of flanking vowels and voicing lasted well into the stop closure and the other had high F_1 or F_0 at flanking vowel edges and voicing stopped early in the stop closure. Stops with opposite combinations of these two properties, low F_1 or F_0 at vowel edge and short-lasting voicing vs. high F_1 or F_0 at vowel edge and long-lasting voicing, were much harder to discriminate, even though the physical size of the differences in both properties was the same for both combinations of values. In the more discriminable pair of stops, energy is either concentrated at low frequencies in and near the stop closure or it isn't (see also Stevens & Blumstein, 1981). The combinations of properties in the less discriminable pair cancel out each other's effects on the concentration of energy at low frequencies.

Closure Duration. Parker, *et al.* (1986), Kingston, *et al.* (1990), and Kingston & Diehl (1995) showed that when the stop closure contains voicing, i.e. low frequency periodic energy, listeners are more likely to label an intervocalic stop with a particular closure duration as [+voice] or short (recall that [+voice] stops have shorter closure durations) than when it contains no voicing. However, they only do so as long as F_1 is low at the flanking vowels' edges (low F_0 at the vowels' edges doesn't have the same effect). Voicing and low F_1 together make the closure more continuous spectrally with the flanking vowels at low frequencies and thereby reduce the perceived interruption of the vowels by the closure.

Consonant:Vowel Duration Ratio. Kluender, *et al.* (1988; see also Kohler (1979), Port & Dalby (1982), and Fowler (1992)) showed that listeners judge a particular closure duration to be [+voice] or short more often if the preceding vowel is long than if it's short. That listeners judge closure duration in contrast to the perceived duration of the preceding vowel suggests that this property is the ratio of these two durations: small values of this ratio induce [+voice] judgments, large values [-voice] judgments.

In each of these cases, the combinations of properties that enhance contrasts or increase [+voice] over [-voice] judgments are those observed in naturally produced stops contrasting for this feature, as laid out in (1). However, listeners' behavior cannot be due entirely their having (over-)learned the combinations of acoustic properties that do occur because all the same perceptual interactions have also been obtained with non-speech analogues of vowel-stop-vowel stimuli, which would not evoke listeners' experience with speech sounds. These interactions or perceptual properties arise from audition not experience.

Moreover, integration's pervasiveness shows that listeners often don't hear the cues as individual properties of speech sounds. Instead, they hear perceptual properties like the low frequency property that arise from the integration of these cues. Picking out one or another cue as essential to conveying a contrast misrepresents the listener's perceptual experience as well as the nature of the perceptual information they find in the signal.

So far, I have shown that the first assumption of the licensing by cue proposal is wrong: that contrasts are pronounced in the same way in all contexts or languages. Rather than being the same, the pronunciations of contrasts differ systematically between contexts and languages. Some of these differences, furthermore, produce combinations of acoustic properties that enhance contrasts. Finally, if speakers can choose different pronunciations in different contexts to improve the distinctiveness of a contrast there, then the contexts themselves don't differ intrinsically in how well contrasts can be realized in them.

1.2. Salience differences between cues?

Do the cues themselves differ in intrinsic perceptual salience? Sound changes in which a contrast is transferred to one of its cues naturally test the relative salience of cues as one would expect the contrast to be transferred to its most salient cue. However, the dis-integration of final [voice] contrasts shows instead that the contrast can be transferred to any of its cues.

The Friulian forms in (2) show the transfer of the [voice] contrast to a vowel duration difference in preceding vowels (Baroni & Vanelli 2000). The stems in (a-d) alternate between a final [+voice] obstruent when followed by a vowel and a [-voice] consonant when word-final. Vowel lengthening accompanies devoicing. The non-alternating stems in (e-h) show that nothing happens when the stem ends in [-voice] obstruent, i.e. when its pronunciation is the same finally as before a vowel. The lengthened vowel before a devoiced obstruent in a form such as ¹*la:t*

Keeping and Losing Contrasts

'gone (m.)' is more than twice as long as the vowel in 'lat 'milk', while one before a [+voice] consonant in a form such as 'lade 'gone (f.)' is only half again longer. The lengthened vowels are also more peripheral, their F_0 peaks occur earlier and F_0 falls more across the vowel, and the following stop closure is shorter than in a comparable non-alternating form. The [voice] contrast has not just been transferred to vowel length but the longer vowel duration that otherwise occurs before [+voice] obstruents in this language has been both greatly exaggerated and augmented by other phonetic differences that enhance it.

(2) [voice] and vowel length alternations in Friulian

a. 'gone (m.)'	'la:t	'lade	'gone (f.)'
b. 'fire'	'fu:k	fogo'lar	'fire place'
c. 'weight'	'pe:s	pe'za	'to weigh'
d. 'snow'	'ne:f	neve'a	'to snow'
e. 'milk'	'lat	la'ta	'to breast-feed'
f. 'piece'	'tɔk	tu'kut	'little piece'
g. 'pass'	'pas	pa'sa	'to pass'
h. 'slap'	pa'taf	pata'fa	'to slap'

In (3) we can see that in the same context where Friulian transferred the [voice] contrast to vowel length, Polish has instead transferred it to vowel height (Steele 1973, Stieber 1973, Gussman 1980, Carlton 1990, Buckley 2001). The stems in (a-d) end in a [+voice] consonant before a vowel, which devoices finally; devoicing is accompanied by raising of the mid vowel [o] to high [u]. The non-alternating stems in (e-h) have a final [-voice] consonant before a vowel as well as finally, and their [o] remains unraised. The transfer in this case reinterprets the low F_1 that would ordinarily precede a [+voice] consonant as a high vowel, which has a lower F_1 value than a mid vowel.

(3) [voice] and [high] alternations in Polish

	Nom. Sg.	Nom. Pl.
a. 'ice'	l[ut]	l[od]y
b. 'corner'	r[uk]	r[og]i
c. 'beans'	b[up]	b[ob]y
d. 'knife'	n[uʃ]	n[oʒ]e
e. 'flight'	l[ot]	l[ot]y
f. 'juice'	s[ok]	s[ok]i
g. 'peasant'	chł[op]	chł[op']y
h. 'basket'	k[of]	k[of]e

This case is complicated in two ways, however. First, vowel height also alternates in stems that end in non-nasal sonorants (4a-e), which don't devoice finally, although not in those that end in nasal sonorants (5a-b), e.g.:

(4) [high] without [voice] alternations in non-nasal sonorant-final Polish stems

a. 'peace'	pok[u]j (gen. sg.)	pok[o]ju (nom. sg.)
b. 'cow'	kr[u]w (gen. pl.)	kr[o]w (nom. sg.)
c. 'moth'	m[u]l (gen. sg.)	m[o]la (nom. sg.)
d. 'barn'	stod[u]ł (gen. pl.)	stod[o]ł (nom. sg.)
e. 'time'	p[u]r (gen. pl.)	p[o]ra (nom. sg.)

(5) Neither [high] nor [voice] alternations in nasal sonorant-final Polish stems

a. 'home'	d[o]m (nom. sg.)	d[o]mu (gen. sg.)
b. 'side'	str[o]n (gen. pl.)	str[o]na (nom. sg.)

Raising before non-nasal sonorants is unsurprising as F_1 is low at the edges of vowels before any voiced consonant, whether it's a sonorant or an obstruent. The failure to raise before nasal sonorants is equally unsurprising as F_1 is instead raised by nasalization, and high nasalized vowels often lower to mid (Maeda, 1993; Hajek, 1997).

The second complication is at first glance more serious, however. (The following discussion is based equally on Stieber (1971) and Buckley (2001).) The alternants in (3-5) where the stem-final consonant is now word-final originally ended in weak yers. These yers were lost with compensatory lengthening of the stem vowel before final voiced consonants (sonorants as well as obstruents) around 1000 CE, e.g. 'ice' *lod-ĭ* (nom. sg.) > *lo:d* vs *lod-y* (nom. pl.). (Present-day forms are used for illustration.) Compensatory lengthening was blocked by a final voiceless consonant, e.g. 'flight' *lot-ĭ* (nom. sg.) > *lot* vs. *lot-y* (nom. pl.). By the late 14th century, the lengthened mid vowels were higher and more peripheral than the unlengthened ones: *lō:d* vs *lōd-y* and 'flight' *lōt* and *lōt-y*. Final obstruents devoiced about the same time or perhaps somewhat earlier in the 14th century: *lō:d* > *lō:t* vs *lōd-y* and 'flight' *lōt* and *lōt-y*. Vowel quantity differences only began to disappear more than a century later, after 1450 CE: *lō:t* > *lōt* vs *lōd-y* and 'flight' *lōt* and *lōt-y*, and later still, certainly after 1600 CE, /o/ and /u/ merged to /u/ and /ɔ/ raised to /o/: *lōt* > *lut* vs *lōd-y* > *lod-y* and 'flight' *lōt* > *lot* and *lōt-y* > *lot-y*. Like Friulian, Polish first develops a longer vowel before a [+voice] consonant, albeit because a [+voice] consonant is compatible with compensatory lengthening, not via transfer of the [voice] contrast. Moreover, raising appears to accompany length.

Nonetheless, the higher vowels emerged before consonants whose voicing would lower F_1 in the preceding vowel, and raising coincided with devoicing in the 14th century. Compensatory lengthening may have been the first step in the

development of the present-day alternation, but the following consonant's voicing remains the most likely cause of raising. By itself, lengthening would instead have lowered vowels, as lower vowels are inherently longer than higher ones. Because word-final obstruents devoiced at the same time, raising immediately came to share the burden of the [voice] contrast with vowel length. Once quantity differences were lost, only vowel height differences remained to convey the contrast. Later mergers led to the present-day high vowel in the bare stem alternants in (3-5).

The last example of dis-integration of a final [voice] contrast underlies a sound change taking place in southern American English. Thomas (2000) and Moreton (2002) show that the offglide in the diphthongs /aɪ, aʊ, ɔɪ, εɪ/ before [-voice] obstruents has lower F₁ values and more extreme F₂ values, higher in the fronting diphthongs /aɪ, ɔɪ, εɪ/ and lower in the backing diphthong /aʊ/, than in other contexts. [Summers (1987) observed more extreme F₁ values and larger jaw movements in /a, æ/ before [-voice] than [+voice] obstruents; both F₁ value and jaw height also changed faster before [-voice] than [+voice] obstruents.] Moreton (2002) also found that listeners identify final stops as [-voice] more often when F₁ is low and F₂ is high in the offglide of /aɪ/. Moreton's speakers and listeners came from northern as well as southern parts of the United States, so more extreme pronunciations of these diphthongs before [-voice] obstruents than elsewhere aren't just found in the south. However, in much of the south, this difference has been phonologized to the extent that /aɪ/ is still pronounced as a diphthong before [-voice] obstruents but has become a low fronted monophthong [a] elsewhere. Although following obstruents contrasting for [voice] usually remain phonetically different, before flaps the difference between a diphthong and monophthong alone conveys the [voice] contrast.

In each of these three examples, what was only one of many phonetic cues to the [voice] contrast, vowel duration, F₁, or diphthongization, is now the only means of distinguishing morphemes that once ended in [+voice] vs. [-voice] consonants. This diversity suggests that none of cues is any more salient than the others, contrary to the second assumption of the licensing by cue proposal that cues differ intrinsically in salience. In short, no cue is privileged by greater perceptual salience.

1.3. Optimal pronunciations

Before ending this critique of licensing by cue, it is useful to take up a case of dis-integration that does not involve the [voice] contrast. Besides showing the generality of dis-integration, this case also shows that speakers exert themselves to convey contrasts in ways that are entirely unexpected if they couldn't optimize their pronunciations to ensure that contrasts are conveyed.

In languages where nasalization is not contrastive in vowels, soft palate height covaries directly with tongue height, causing lower vowels to be more nasalized than higher vowels (see Kingston (1991) for a review of the evidence). English is such a language, and for English listeners nasalization separates perceptually from

height in higher vowels, but integrates with height in lower vowels: greater acoustic nasalization makes a lower vowel sound lower rather than more nasalized (Kingston & Macmillan 1995, Macmillan *et al.* 1999). More nasal coupling is needed to get a lower vowel to sound nasalized (Maeda, 1993).

The perceptual integration of nasalization into the percept of vowel lowness and the need for more nasal coupling are unexpected given the ubiquity of low vowels in nasal vowel inventories. In the expanded version of UPSID with 451 languages (Maddieson & Precoda 1992) are 101 that contrast nasal with oral vowels. 49 of these languages have the same number of nasal as oral vowels, but in the remaining 52, one or more nasal vowels corresponding to oral vowels are missing. The examples in (6) show the patterns and frequencies of gaps in the nasal vowel inventories of these languages:

(6) Typology of gaps in inventories of nasal vowels corresponding to oral vowels

a. *Headless*, missing high nasal vowels (21 languages), e.g., Amuzgo:

Oral		Nasal	
i	u		
e	o	ẽ	õ
æ	a	ã	ã

b. *Gutless*, missing mid nasal vowels (39 languages), e.g., Senadi:

Oral		Nasal	
i	u	ĩ	ũ
ɛ	ɔ		
e	o	ẽ	õ
a		ã	

c. *Footless*, missing low nasal vowels (6 languages), e.g., Chatino:

Oral		Nasal	
i	u	ĩ	ũ
e	o	ẽ	õ
a			

The gutless type (6b), where mid nasal vowels are missing, outnumbers the headless type (6a), where high nasal vowels are missing, by nearly 2:1. The footless type (6c) is comparatively very rare. How can nasal vowel inventories nearly always include a low nasal vowel if nasalization is integrated into the percept of vowel lowness in lower vowels?

Speakers must lower the soft palate more in lower than higher vowels to ensure that they're perceived as nasal as well as low. Fortunately, the vowel also sounds lower because some of that nasalization is still integrated into the percept of vowel lowness. Low vowels' inherently greater duration may also help make

the spectral modifications caused by nasalization easier to detect (Hajek 1997; see also Krakow, *et al.* (1988) and Whalen & Beddor (1989)).

Otherwise, gutless nasal vowel inventories are most common because mid nasal vowels either lower, if nasalization is perceived as lowness, or raise, to improve the perceptual separability of nasalization. Headless nasal vowel inventories are next most common because high vowels are inherently short and may thus sometimes be too brief for nasalization to be perceived reliably. Nasalization either doesn't become contrastive in high vowels or integration lowers them to mid.

In this case, the most common instance of a type, a low nasal vowel, is the one that requires the most extreme articulation, the greatest soft palate lowering, to ensure that its acoustic effect is dis-integrated perceptually from other properties of that sound. Speakers must be altruists.

2. What are releases good for?

2.1. Do releases package acoustic information about segment identity?

Much of the concrete discussion of licensing by cue has turned on the presence and perceptual value of an audible release in consonants. The release's acoustics are thought to bear much of the burden of conveying the consonants' phonation, place, and manner contrasts. I bear no little responsibility for this emphasis, having argued in Kingston (1985, 1990) that releases had these virtues. Here, I first review the evidence and arguments that led me and others to emphasize the importance of releases for segment identification and then turn to the perceptual evidence which shows that listeners don't after all rely on the acoustics of consonant releases for this purpose. Finally, I review evidence which suggests that releases, along with much other allophonic variation, are probably perceptually valuable because they aid listeners in segmenting the utterance prosodically, which in turn helps them find words in the stream of speech.

My original argument was that the release of a consonant constriction "packages" the consonant's values for phonation, manner, and place of articulation into a brief, salient acoustic event (see also Stevens & Keyser (1989) and Liu (1996)). This event might be salient because neurons in the VIIIth (auditory) nerve fire robustly for a brief time when signal energy rises abruptly; subsequently, their firing rate quickly drops off (Delgutte & Kiang 1984a,b; Delgutte 1996; Silverman 1997; Wright 1999, 2001). The representation of spectral energy distributions in the peripheral auditory system as well as higher up should therefore be best at the release, particularly of a non-continuant consonant.

Stevens & Keyser (1989) showed how the most common contrasts for manner and place of articulation in the world's languages correspond to large differences in how spectral energy distributions change at consonant releases. Earlier, Stevens & Blumstein (1978; Blumstein & Stevens 1979, 1980) had shown that stops could be automatically classified for place of articulation using just gross features of spectra calculated across a short interval (25.6 ms) beginning with the release, and Kewley-Port, *et al.* (1983) and Lahiri, *et al.* (1984) had been able to classify

stops' place of articulation automatically using changes in the spectral distribution of energy across a brief interval between the release and a point early in the following vowel. Released final stops are also identified much better than unreleased ones (Malécot, 1958), and in a cluster of stops between vowels, only the released one is heard if the closure duration is short enough that the listener thinks only a single stop has been pronounced (Fujimura, *et al.* 1978; Repp 1978, 1983; Ohala 1990).

Nonetheless, releases don't convey the most salient information for either place or manner perception. When release spectra are pitted against formant transitions in place perception, formant transitions nearly always determine what place of articulation is perceived (Walley & Carrell 1983, Smits *et al.* 1996a,b). The rate at which energy rises at signal onset, a.k.a. "rise time", should directly affect VIIIth nerve firing rate, yet rise time contributes little or nothing to perceiving the [continuant] contrast between affricates vs. fricatives—listeners rely on noise duration instead (Kluender & Walsh 1992)—nor to the combined [continuant, sonorant] contrast between stops vs. glides—listeners rely on transition duration instead (Diehl & Walsh 1989, Walsh 1991).

The original idea that releases packaged segmental information assumed, like the licensing by cue proposal it inspired, that it wasn't sounds' pronunciations that varied systematically with context but instead only the audibility of cues to their identity. Consonants are always released, but the release is only audible in some contexts, specifically when the following sound has a much more open articulation than the consonant, ideally a vowel or a non-nasal sonorant. Before such a more open articulation, energy would be audible across the spectrum, and any effects of glottal or oral articulations on its distribution could be detected. As consonants preceding a vowel or non-nasal sonorant are most likely to be syllabified into the onset with that sound, releases are much more likely to be audible and informative about segment identity in syllable onsets than codas (Kingston 1985; Lombardi 1991, 1995).

However, speakers actually articulate consonants differently in onsets than they do in codas. Soft palate and lip movements are closely synchronized in [m] in onsets, but soft palate movement begins long before lip movement in [m] in codas (Krakow 1989, 1993). Similarly, tongue dorsum and tip movements are closely synchronized in [l]s in onsets, but tongue dorsum movement begins long before tip movement in [l]s in codas (Sproat & Fujimura 1983). Kelso, *et al.* (1986) had speakers repeat either [ip] or [pi] and then speed up repeatedly. In repetitions of [pi], peak glottal opening was delayed relative to lip opening onset by a constant proportion at all speaking rates, but in repetitions of [ip], peak opening abruptly shifted from coinciding with lip opening onset to a delay equal to that observed in [pi] when the speaking rate exceeded about 4 syllables/second. Although this shift was apparently involuntary, it shows that glottal opening is timed differently relative to lip opening for a voiceless stop in a coda from one in an onset. Tuller & Kelso (1991) show that these differences affect listeners' percept of the stop's syllabification. Finally, Smith (2002) shows that the

pronunciation of segments in onsets and other *prominent* positions may even neutralize contrasts in order to increase the position's salience. For example, in Chamicuro, the glottal consonants /ʔ, h/ are prohibited from occurring in syllable onsets; that is, the contrast between consonants with and without an oral constriction is neutralized in favor of those with an oral constriction in this position. These results all show that speakers strive to pronounce onsets differently from codas; consonants don't merely sound different in the two syllable positions.

2.2. Do releases signal prosodic boundaries instead?

Even if listeners don't rely on the acoustic information packaged in releases to identify segments, do they use releases for any other purpose? Because a consonant is usually only released audibly when it precedes a segment with a more open articulation and in that case is usually syllabified with that segment, the release may convey that syllabification to the listener. The absence of an audible release would instead convey that the consonant is not syllabified with the following segment. If this hypothesis is correct, then releases are simply one of a large number of allophonic properties that convey the grouping of segments in syllables, as well into higher-level constituents in the prosodic hierarchy.

There is also considerable evidence that the allophones which occur at the beginning of prosodic constituents interrupt the signal more than those which occur inside prosodic constituents (English: Pierrehumbert & Talkin 1992, Turk 1993, Dille, *et al.* 1996; French: Fougeron & Keating 1997, Fougeron 2000; and Korean: Silva 1992, Jun 1993, Cho & Keating 2000). A greater interruption separates the next prosodic constituent more from the preceding one. An audible release may not only link a consonant prosodically to the following segment, but also increase the interruption of the signal and thereby sharpen the separation of prosodic constituents.

Although the listener would benefit from the marking of the beginning of any prosodic constituent by releases and other allophones, these phonetic events could be particularly useful for finding the beginnings of phonological words in the stream of speech. Initial allophones which aren't word-initial inhibit word spotting in longer non-word strings in Dutch (McQueen 1998) and English (Kirk 2000, 2001). For example, *wine* is harder to spot following the aspirated [k^h] in [vuk^hwain] than the unaspirated [k] in [vukwain] and *rock* is harder to spot following the retroflexed, affricated [dʒ] in [vudʒak] than the alveolar stop [d] in [vudjak]. Kirk argued that the initial allophones [k^h] and [dʒ] lead listeners to syllabify the end of the preceding residue with the target word, and the resulting mismatch between syllable and word boundaries makes the word harder to spot than when the residue ends instead with a non-initial allophone. Both allophonic variations are furthermore differences in the presence vs absence of a particular kind of release of the final consonant of the residue.

2.3. Summary

In this section of the paper, I have tried to show that releases aren't packages of acoustic information about segment identity, at least not ones that listeners rely on much. Instead, releases are one of a large number of allophonic variants that signal the beginning of prosodic constituents and thus aid the listener in parsing the stream of speech into words and higher level prosodic constituents. Word-spotting experiments show that listeners do indeed use releases for this purpose.

3. Listening in context

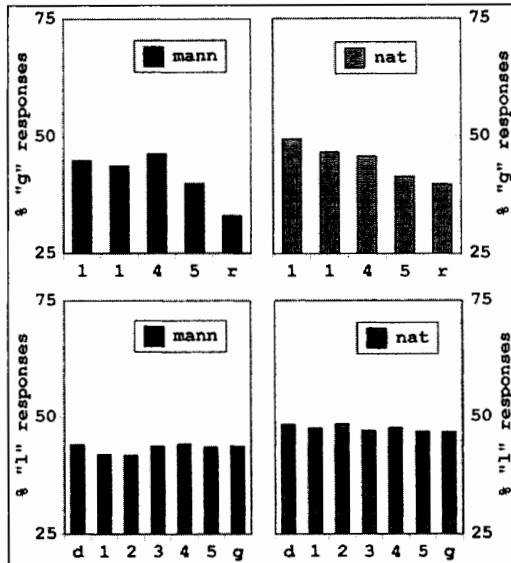
3.1. Perceptual contrast between adjacent segments

It is a truism that in everyday life, speech sounds are pronounced and must be recognized next to other speech sounds. Both a very general and a very specific characteristic of neighboring sounds' effects on the perception of the current sound phonetically explain the distribution of contrast and neutralization and thus replace the licensing by cue account of this distribution that I argued against in §1 above.

Nearey (1990, 1995, 1997, in press), Smits (2001a,b), and Benkí (this volume) have shown that listeners extract segments from the speech stream even though perceptually relevant information often comes from acoustic intervals corresponding to adjacent segments and neighboring segments mutually influence one another's recognition. The effect of a neighboring segment is typically contrastive: next to a segment which is at one extreme of some acoustic dimension, a segment which is intermediate along that dimension will sound like it's at the opposite extreme. For example, an intermediate member of a [d-g] continuum synthesized by incrementally varying F_3 onset frequency is more likely to be labeled heard as the continuum's low endpoint "g" after [l], whose F_3 is high, than after [ɹ], whose F_3 is low (Mann 1980). Lotto & Kluender (1998) obtained the same "g" bias when the liquid was replaced by a pure tone whose frequency corresponded to the liquid's F_3 , which shows the contrast arises from audition rather than compensation for coarticulation (cf. Fowler, *et al.* 1990, 1999). Gouskova & Kingston (2002, unpublished data) recently replicated Mann's results using two liquid continua as contexts, one synthesized with Mann's parameter values (*mann* stimuli) and the other with values based on Stevens's 1998 description and modeling of naturally produced liquids' acoustics (*nat* stimuli). Listeners identified the liquid as well as the stop. The top two panels in Figure (7) show that "g" responses drop by about 10% as the liquid goes from [l] to [ɹ], and the bottom two panels show that the percentage of "l" responses changes little as the stop goes from [d] to [g]. The effects and their sizes don't differ between the *mann* and the *nat* stimuli. The progressive effect of the liquid on stop judgments is contrastive, but there is no regressive effect of the stop on liquid judgments.

Keeping and Losing Contrasts

- (7) Frequency of “g” responses as a function of the preceding liquid (top) and “l” responses as a function of the following stop (bottom), for *mann* (left) vs *nat* (right) liquids.



Other context effects are also contrastive. An intermediate member of a [t-k] continuum, also synthesized by varying F_3 onset frequency, is more likely to be labeled the high endpoint “t”, after [ʃ], whose noise is concentrated at relatively low frequencies, than after [s], whose noise is concentrated at relatively high frequencies (Mann & Repp 1981; Repp & Mann 1981, 1982). Consonants’ place also contrastively shifts the labeling of neighboring vowels’ backness. Listeners label intermediate members of an [I-U] continuum more often as the high F_2 endpoint “i” between labiovelar glides [w_w], whose F_2 values are low, than between palatal glides [j_j], whose F_2 values are high. Similarly, listeners label intermediate members of a steady-state [i-u] continuum more often as the high F_2 endpoint “i” between the labials [f_p], whose energy is concentrated at relatively low frequencies, than between the coronals [s_t], whose energy is instead concentrated at high frequencies (Ohala *et al.* 1978). Listeners also label a steady-state [i-u] continuum’s intermediate members as “i” more often before labial [b] than coronal [d], whether they actually hear the stop or restore it (Ohala & Feder 1995; Bradlow & Kingston 1990). Finally, Holt *et al.* (2000) report that listeners label intermediate members of an [ε-Λ] continuum more often as “ε” between labials [b_b] than coronals [d_d], as well as FM glides mimicking the F_2 and F_3 transitions to and from [b] vs. [d] and pure tones matching the F_2 and F_3 onset frequencies next to [b] vs. [d]. In all these cases, the listeners is more likely to

respond with the category corresponding to the continuum's high frequency endpoint when the neighboring segment concentrates energy at relatively low frequencies than when it instead concentrates it at relatively high frequencies.

3.2. Perceptual assimilation as well as contrast

Although context effects are nearly always contrastive, like these examples, in one circumstance they are consistently assimilatory. In stop-stop clusters between vowels (VC_1C_2V ; Repp 1978, 1983; Ohala 1990), C_1 is perceived as having the same place of articulation as C_2 when the closure duration is short enough (relative to the range of closure durations heard) to be just one consonant rather than two. Alternatively, the listener hears only C_2 when the closure duration is short. Even if C_1 isn't heard as a separate segment at these short closure durations, its presence and distinct place of articulation nonetheless contrastively shifts C_2 labeling. At closure durations long enough to be two consonants rather than just one, labeling is mutually contrastive, although the contrastive effect of C_1 on C_2 remains stronger than the reverse. Shinya & Kingston (2002, unpublished data) replicated Repp (1983) using a 7-step [b-d-g] continuum synthesized in both VC_1 and C_2V by varying only F_2 and F_3 offset and onset frequencies. VC_1 and C_2V were mirror images in that no burst preceded formant onset in C_2V . VC_1 and C_2V were separated by silent gaps lasting 50, 100, 150, or 200 ms. Listeners labeled the stops as "b", "d", or "g" if they heard just one, and labeled both C_1 and C_2 as one of these three alternatives if they instead heard two. Assimilation and contrast were quantified using the formulae like those in (8) (following Repp 1983):

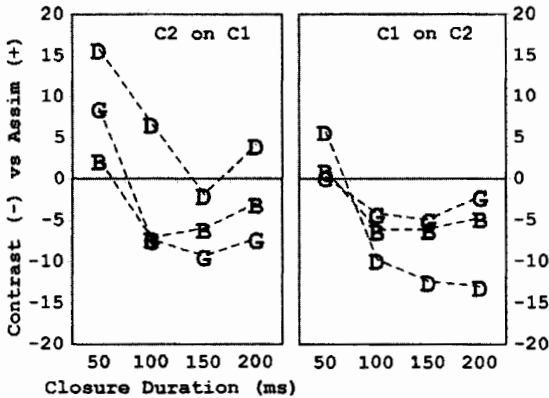
- (8) a. Assimilation: $[\% \text{ "b" } / V_bV] - [\% \text{ "b" } / V_ \{d,g\} V] > 0$
 b. Contrast: $[\% \text{ "b" } / V_bV] - [\% \text{ "b" } / V_ \{d,g\} V] < 0$

These two formulae assess the regressive effect of C_2 on the entire range of C_1 stimuli: if more "b" responses are given before [b] than before [d] or [g], then C_1 has assimilated to C_2 , but if "b" responses before [d] or [g] instead outnumber those before [b], then C_1 instead contrasts with C_2 . Scores were calculated for "b" responses before the best [b] vs. the best [d] and [g], and "b" responses before the best [b] stimulus vs. all other stimuli. The resulting scores differed very little, so only those calculated the second way are presented here. Entirely parallel calculations were carried out for assessing the progressive effect of C_1 on C_2 and for the other places of articulation. These scores are displayed as a function of closure duration for the regressive effects of C_2 on C_1 in the left panel of Figure (9), and for the progressive effects of C_1 on C_2 in the right panel. At the shortest closure duration, a listener is more than 15% more likely to label C_1 as "d" when C_2 is the best [d] exemplar than when C_2 is any other stimulus, and a bit less than 10% more likely to label C_1 as "g" when C_2 is the best [g] exemplar than when C_2 is any other stimulus. "b" responses to C_1 are at most just slightly more likely before the best [b] exemplar than other stimuli at this closure duration. At all longer closure durations, both "b" and "g" responses to C_1 are instead more likely

Keeping and Losing Contrasts

(5-10%) when C_2 is any other stimulus than the best [b] and [g] exemplars, respectively. “d” responses, on the other hand, remain more likely before the best [d] exemplar than other stimuli for all closure durations except 150 ms, where “d” responses are equally likely before all C_2 s.

(9) Regressive (left) and progressive (right) assimilation (+ %) vs. contrast (- %) in VC1C2V for “b”, “d”, and “g” responses as a function of closure duration.



Listeners are somewhat more likely (5%) to label C_2 as “d” following the best [d] exemplar at the shortest closure duration, but at all other closure durations, they are much more likely (up to 15%) to label C_2 as “d” following any other stimulus. Neither “b” nor “g” responses differ as a function of C_1 at the shortest closure duration and both become somewhat more likely (5%) following any stimulus other than the best exemplars of [b] and [g], respectively, at longer closure durations.

Coronal C_1 percepts are assimilative or neutral across closure durations, but non-coronal C_1 percepts are instead contrastive, except for dorsal C_1 percepts at the shortest duration. C_2 both weakens the percept that C_1 has occurred and that C_1 is different from C_2 when C_1 is perceived to have occurred. Both effects are stronger for coronals than non-coronals. The phonological consequences of this weakening are that C_1 is likely to delete, assimilate to C_2 , or neutralize for place of articulation, especially if C_1 is coronal. On the other hand, coronal C_2 percepts are neutral or strongly contrastive, and non-coronal C_2 percepts are neutral or weakly contrastive. C_1 strengthens the percept that C_2 has occurred and that it’s different from C_1 . Again, both effects are stronger for coronals than non-coronals. C_2 is therefore unlikely to delete, assimilate to C_1 , or neutralize for place, particularly if it’s a coronal. These perceptual asymmetries explain both the much greater susceptibility of coda than onset consonants to loss, assimilation, and neutralization and the particular susceptibility of coronal consonants in codas to these losses of contrast.

another and no burst preceded C₂V. It is only the position of the consonants in the coda that puts them at perceptual disadvantage, not their lack of any place information or acoustic substance that the consonants in the onset possess.

These preliminary results are very encouraging but they leave a number of questions unanswered. First and most basic, why are progressive effects contrastive while regressive effects are assimilative? The onset consonant may dominate simply because it's more recent, but this hypothesis hasn't been tested yet. Second, would similar perceptual asymmetries be observed between onsets and codas for manner and phonation contrasts, which are also lost in codas but kept in onsets? The necessary experiments simply haven't been done. Third, how can this explanation of the distribution of contrast and neutralization be generalized to word-final position, where contrasts are also often lost? In texts from 20 disparate languages, Janda (1979) found that a word is far more likely to be followed by a word beginning with a consonant than one beginning with a vowel. He argued that a following word boundary and a following consonant are so often common environments for phonological constraints or processes because # is in fact usually C. The conditions for perceptual assimilation and contrast between coda and onset consonants therefore usually exist at the ends of words as well as inside them.

4. Summary

In this paper, I have tried to show three things. First, contrasts aren't licensed by cue but instead by the adjustments and effort speakers make to ensure their differences are perceived. Second, listeners don't rely on releases as packages of salient acoustic information for identifying segments but instead for segmenting the signal prosodically, particularly into prosodic words. Third, a sound is most often heard as different from its neighbors, except when it's a consonant before another consonant, in which case it's instead likely to be heard as the same as the following consonant, while the following consonant is likely to be heard as different from the preceding one. This perceptual asymmetry may explain why contrasts are typically lost in codas but kept in onsets. As nearly all the neutralizations that licensing by cue purports to explain occur in codas, this alternative explanation covers the ground equally well. Furthermore, it depends on a quite general and well-established perceptual interaction between neighboring segments.

References

- Baroni, M., & Vanelli, L. 2000. The relationship between vowel length and consonantal voicing in Friulian, In L. Repetti (ed.), *Phonological Theory and the Dialects of Italy*, 13-44. Amsterdam, Philadelphia: John Benjamins Pub. Co.

Keeping and Losing Contrasts

- Benkí, J. R. this volume. Effects of signal-independent factors in speech perception, *Proc. 28th Ann. Mtg. Berkeley Ling. Soc.*
- Blumstein, S. E., & Stevens, K. N. 1979. Acoustic invariance in speech production: Evidence from the measurement of the spectral characteristics of stop consonants, *J. Acoust. Soc. Am.*, **66**: 1001-1017.
- Blumstein, S. E., & Stevens, K. N. 1980. Perceptual invariance and onset spectra for stop consonants in different vowel environments, *J. Acoust. Soc. Am.*, **67**: 648-662.
- Bradlow, A. R., & Kingston, J. 1990. Cognitive processing in the perception of the speech. *J. Acoust. Soc. Am.*, **88**: S56. (Abstract).
- Buckley, E. 2001. Polish o-raising and phonological explanation, *Ann. Mtg. Ling. Soc. Am.*, Washington, DC.
- Carlton, T. R. 1990. *Introduction to the Phonological History of the Slavic Languages*, Columbus: Slavica.
- Chen, M. 1970. Vowel length variation as a function of the voicing of the consonant environment, *Phonetica*, **22**: 125-159.
- Cho, T., & Keating, P. 2000. Articulatory and acoustic studies on domain-initial strengthening in Korean, *J. Phon.*, **29**: 155-190.
- Davis, S., & Summers, W. V. 1989. Vowel length and closure duration in word-medial VC sequences, *J. Phon.*, **17**: 339-353.
- Delgutte, B., & Kiang, N. Y. S. 1984a. Speech coding in the auditory nerve III: Voiceless fricative consonants, *J. Acoust. Soc. Am.*, **75**: 887-897.
- Delgutte, B., & Kiang, N. Y. S. 1984b. Speech coding in the auditory nerve IV: Sounds with consonant-like dynamic characteristics, *J. Acoust. Soc. Am.*, **75**: 897-907.
- Delgutte, B. 1996. Auditory neural processing of speech. In W. J. Hardcastle and J. Laver (eds.), *Handbook of Phonetic Sciences*, 507-538. Oxford: Blackwell Publishing.
- Diehl, R. L., & Walsh, M. A. 1989. An auditory basis for the stimulus-length effect in the perception of stops and glides, *J. Acoust. Soc. Am.*, **85**: 2154-2164.
- Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. 1996. Glottalization of word-initial vowels as a function of prosodic structure, *J. Phon.*, **24**: 423-444.
- Fougeron, C. 2000. Articulatory properties of initial segments in prosodic constituents in French, *J. Phon.*, **29**: 109-135.
- Fougeron, C., & Keating, P. A. 1997. Articulatory strengthening at edges of prosodic domains, *J. Acoust. Soc. Am.*, **106**: 3728-3740.
- Fowler, C. A. 1992. Vowel duration and closure duration in voiced and unvoiced stops: There are no contrast effects here, *J. Phon.*, **20**: 143-166.
- Fowler, C. A., Best, C. T., & McRoberts, G. W. 1990. Young infants' perception of liquid coarticulatory influences on following stop consonants, *Percept. & Psychophys.*, **48**: 559-570.

- Fowler, C. A., Brown, J. M., & Mann, V. A. 1999. Compensation for coarticulation in audiovisual speech perception, *Proc. XIVth Intl. Cong. Phon. Sci.*, San Francisco, 639-642.
- Fujimura, O., Macchi, M. J., & Streeter, L. A. 1978. Perception of stop consonants with conflicting transitional cues: A cross-linguistic study, *Lg. & Sp.*, **21**: 337-346.
- Gussman, E. 1980. *Studies in Abstract Phonology*, Cambridge, MA: MIT Press.
- Hajek, J. 1997. *Universals of Sound Change in Nasalization*, Publications of the Philological Society, 31, Oxford, UK: Blackwell Publishers.
- Holt, L. R., Lotto, A. J., & Kluender, K. R. 2000. Neighboring spectral context influences vowel identification, *J. Acoust. Soc. Am.*, **108**: 710-722.
- Janda, R. 1979. Double-cross in phonology: Why word boundary (often) acts like a consonant, *Proc. 5th Ann. Mtg. Berkeley Ling. Soc.*, 397-412.
- Jun, S.-A. 1993. *The Phonetics and Phonology of Korean Prosody*, Ph.D. dissertation, Ohio State University.
- Keating, P. A. 1984. Phonetic and phonological representation of stop consonant voicing, *Lg.*, **60**: 286-319.
- Kelso, J. A. S., Saltzman, E. L., & Tuller, B. 1986. The dynamical perspective on speech production: Data and theory, *J. Phon.*, **14**: 29-59.
- Kewley-Port, D., Pisoni, D. B., & Studdert-Kennedy, M. 1983. Perception of static and dynamic acoustic cues to place of articulation in initial stop consonants, *J. Acoust. Soc. Am.*, **73**: 1779-1793.
- Kingston, J. 1985. *The Phonetics and Phonology of the Timing of Oral and Glottal Events*, Ph.D. dissertation, University of California, Berkeley.
- Kingston, J. 1990. Articulatory binding, In J. Kingston & M. E. Beckman (eds.), *Papers in Laboratory Phonology I*, 406-434. Cambridge, UK: Cambridge University Press.
- Kingston, J. 1991. Integrating articulations in the perception of vowel height, *Phonetica*, **47**: 149-179.
- Kingston, J., & Diehl, R. L. 1994. Phonetic knowledge, *Lg.*, **70**: 419-454.
- Kingston, J., & Diehl, R. L. 1995. Intermediate properties in the perception of distinctive feature values, In A. Arvaniti & B. Connell (eds.), *Papers in Laboratory Phonology IV*, 7-27. Cambridge, UK: Cambridge University Press.
- Kingston, J., Diehl, R. L., Kirk, C. J., & Castleman, W. A. submitted. On the internal perceptual structure of phonological features: the [voice] distinction, Ms. University of Massachusetts, Amherst and University of Texas, Austin.
- Kingston, J., & Macmillan, N. A. 1995. Integrality of nasalization and F_1 in vowels in isolation and before oral and nasal consonants: A detection-theoretic application of the Garner paradigm, *J. Acoust. Soc. Am.*, **97**: 1261-1285.
- Kirk, C. J. 2000. Syllabic cues to word segmentation, *Proc. Workshop Spoken Word Access Processes*, 131-134. Nijmegen: Max Planck Inst. Psycholing.

- Kirk, C. J. 2001. *Phonological Constraints on the Segmentation of Continuous Speech*, Ph.D. dissertation, U. Massachusetts, Amherst.
- Kohler, K. J. 1979. Dimensions in the perception of fortis and lenis plosives, *Phonetica*, **36**: 332-343.
- Kluender, K. R., Diehl, R. L., & Wright, B. A. 1988. Vowel-length differences before voiced and voiceless consonants: An auditory explanation, *J. Phon.*, **16**: 153-169.
- Kluender, K. R., & Walsh, M. A. 1992. Amplitude rise time and the perception of the voiceless affricate-fricative distinction, *Percept. & Psychophys.*, **51**: 328-333.
- Krakow, R. A. 1989. *The Articulatory Organization of Syllables: A Kinematic Analysis of Labial and Velar Gestures*, Ph.D. dissertation, Yale University.
- Krakow, R. A. 1993. Nonsegmental influences on velum movement patterns: Syllables, sentences, stress, and speaking rate, In M. K. Huffman & R. A. Krakow (eds.), *Nasals, Nasalization, and the Velum, Phonetics and Phonology*, **5**, 87-116. San Diego: Academic Press.
- Krakow, R. A., Beddor, P. S., Goldstein, L. M., & Fowler, C. A. 1988. Coarticulatory influences on the perceived height of nasal vowels, *J. Acoust. Soc. Am.*, **83**: 1146-1158.
- Laeuffer, C. 1992. Patterns of voicing-conditioned vowel duration in French and English, *J. Phon.*, **20**: 411-440.
- Lahiri, A., Gewirth, L., & Blumstein, S. E., 1984. A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: Evidence from a cross-language study, *J. Acoust. Soc. Am.*, **76**: 391-404.
- Lindblom, B., & Studdert-Kennedy, M. 1967. On the role of formant transitions in vowel recognition, *J. Acoust. Soc. Am.*, **42**: 830-843.
- Lisker, L. 1986. "Voicing" in English: A catalogue of acoustic features signaling /b/ vs /p/ in trochees. *Lg. & Sp.*, **29**: 3-11.
- Liu, S. A. 1996. Landmark detection for distinctive feature-based speech recognition, *J. Acoust. Soc. Am.*, **100**: 3417-3430.
- Lombardi, L. 1991. *Laryngeal Features and Laryngeal Neutralization*, Ph.D. dissertation, University of Massachusetts, Amherst.
- Lombardi, L. 1995. Laryngeal neutralization and syllable wellformedness, *Nat. Lg. Ling. Theory*, **13**: 39-74.
- Lotto, A. J., & Kluender, K. R. 1998. General contrast effects in speech perception: Effect of preceding liquid on stop consonant identification, *Percept. & Psychophys.*, **60**: 602-619.
- Mack, M. 1982. Voicing-dependent vowel duration in English and French: Monolingual and bilingual production, *J. Acoust. Soc. Am.*, **71**: 173-178.
- Macmillan, N. A., Kingston, J., Thorburn, R., Walsh Dickey, L., & Bartels, C. 1999. Integrality of nasalization and F₁. II. Basic sensitivity and phonetic labeling measure distinct sensory and decision-rule interactions, *J. Acoust. Soc. Am.*, **106**: 2913-2932.

- Maddieson, I. this volume. Phonetics in the field, Proc. 28th Ann. Mtg. Berkeley Ling. Soc.
- Maddieson, I., & Precoda, K. 1992. *UPSID and Phoneme, User Manual*, Los Angeles, University of California.
- Maeda, S. 1993. Acoustics of vowel nasalization and articulatory shifts in French nasal vowels, In M. K. Huffman & R. A. Krakow (eds.), *Nasals, Nasalization, and the Velum, Phonetics and Phonology 5*, 147-167. San Diego: Academic Press.
- Malécot, A. 1958. The role of releases in the identification of released final stops, *Lg.*, **34**: 370-380.
- Mann, V. A. 1980. Influence of preceding liquid on stop-consonant perception, *Percept. & Psychophys.*, **28**: 407-412.
- Mann, V. A., & Repp, B. H. 1981. Influence of preceding fricative on stop consonant perception, *J. Acoust. Soc. Am.*, **69**, 548-558.
- McQueen, J. M. 1998. Segmentation of continuous speech using phonotactics, *J. Mem. Lg.*, **39**: 21-46.
- Moreton, E. 2002. Realization of the English postvocalic [voice] contrast in F₁ and F₂, Ms., Johns Hopkins University.
- Nearey, T. M. 1990. The segment as a unit of speech perception, *J. Phon.*, 347-373.
- Nearey, T. M. 1995. A double-weak view of trading relations: Comments on Kingston and Diehl, In A. Arvaniti & B. Connell (eds.), *Papers in Laboratory Phonology IV*, 28-40. Cambridge, UK: Cambridge University Press.
- Nearey, T. M. 1997. Speech recognition as pattern recognition, *J. Acoust. Soc. Am.*, **101**, 3241-3253.
- Nearey, T. M. in press. The factorability of phonological units in speech perception: Simulating the results of speech reception in noise, In *Papers in Laboratory Phonology VI*, Cambridge, UK: Cambridge University Press.
- Ohala, J. J. 1990. The phonetics and phonology of aspects of assimilation, In J. Kingston & M. E. Beckman (eds.), *Papers in Laboratory Phonology I*, 258-275. Cambridge, UK: Cambridge University Press.
- Ohala, J. J., & Feder, D. 1995. Listeners' identification of speech sounds is influenced by adjacent 'restored' phonemes, *Phonetica*, **51**: 111-118, 1994.
- Ohala, J. J., Riordan, C. J., & Kawasaki, H. 1978. The influence of consonant environment upon identification of transitionless vowels, *J. Acoust. Soc. Am.*, **64**: S18. (Abstract)
- Parker, E. M., Diehl, R. L., Kluender, K. R. 1986. Trading relations in speech and non-speech, *Percept. & Psychophys.*, **39**: 129-142.
- Pierrehumbert, J. B. & Talkin, D. 1992. Lenition of /h/ and glottal stop, In G. R. Docherty & D. R. Ladd (eds.), *Papers in Laboratory Phonology II*, 90-116. Cambridge, UK: Cambridge University Press.
- Port, R., & Dalby, J. 1982. Consonant/vowel ratio as a cue for voicing in English, *Percept. & Psychophys.*, **32**: 141-152.

Keeping and Losing Contrasts

- Repp, B. H. 1978. Perceptual integration and differentiation of spectral information of spectral cues for intervocalic stop consonants, *Percept. & Psychophys.*, **24**: 471-485.
- Repp, B. H. 1983. Bidirectional context effects in the perception of VC-CV sequences, *Percept. & Psychophys.*, **33**: 147-155.
- Repp, B. H., & Mann, V. A. 1981. Perceptual assessment of fricative-stop coarticulation, *J. Acoust. Soc. Am.*, **69**: 1154-1163.
- Repp, B. H., & Mann, V. A. 1982. Fricative-stop coarticulation: Acoustic and perceptual evidence, *J. Acoust. Soc. Am.*, **71**: 1562-1567.
- Silva, D. J. 1992. *The Phonetics and Phonology of Stop Lenition in Korean*, Ph.D. dissertation, Cornell University.
- Silverman, D. 1995. *Phasing and Recoverability*, Ph.D. dissertation, U. California, Los Angeles.
- Smith, J. 2002. *Phonological Augmentation in Prominent Positions*, Ph.D. dissertation, U. Massachusetts, Amherst.
- Smits, R. 2001a. Evidence for hierarchical categorization of coarticulated phonemes, *J. Exptl. Psych.: Human Percept. Perform.*, **27**: 1145-1162.
- Smits, R. 2001b. Hierarchical categorization of coarticulated phonemes: A theoretical analysis, *Percept. & Psychophys.*, **63**: 1109-1139.
- Smits, R., ten Bosch, L., & Collier, R. 1996a. Evaluation of various sets of acoustic cues for the perception of prevocalic stop consonants. I. Perception experiment, *J. Acoust. Soc. Am.*, **100**: 3852-3864.
- Smits, R., ten Bosch, L., & Collier, R. 1996b. Evaluation of various sets of acoustic cues for the perception of prevocalic stop consonants. II. Modeling and evaluation, *J. Acoust. Soc. Am.*, **100**: 3865-3881.
- Sproat, R., & Fujimura, O. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation, *J. Phon.*, **21**: 291-311.
- Steele, R. D. 1973. *The Segmental Phonology of Contemporary Standard Polish*, Ph.D. dissertation, Harvard University.
- Steriade, D. 1995. Positional neutralization, Ms. University of California, Los Angeles.
- Steriade, D. 2000. Directional asymmetries in place perception, In E. Hume & K. Johnson (eds.), *The Role of Speech Perception in Phonology*, 219-250. San Diego: Academic Press.
- Steriade, D. 2001. The phonology of perceptibility effects: The P-Map and its consequences for constraint organization, Ms., University of California, Los Angeles.
- Stevens, K. N. 1998. *Acoustic Phonetics*, Cambridge, MA: MIT Press.
- Stevens, K. N., & Blumstein, S. E. 1978. Invariant cues for place of articulation in stop consonants, *J. Acoust. Soc. Am.*, **64**: 1358-1368.
- Stevens, K. N., & Blumstein, S. E. 1981. The search for invariant acoustic correlates of phonetic features, In P. D. Eimas & J. L. Miller (eds.), *Perspectives on the Study of Speech*, 1-38. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Stevens, K. N., & Keyser, S. J. 1989. Primary features and their enhancement in consonants, *Lg.*, **65**: 81-106.
- Stieber, Z. 1973. *A Historical Phonology of the Polish Language*, Heidelberg: Carl Winter Universitätsverlag.
- Summers, W. V. 1987. Effect of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses, *J. Acoust. Soc. Am.*, **82**: 847-863.
- Thomas, E. R. 2000. Spectral differences in /ai/ offsets conditioned by the voicing of the following consonant, *J. Phon.*, **28**: 1-25.
- Tuller, B., & Kelso, J. A. S. 1991. Production and perception of syllable structure, *J. Sp. Hear. Res.*, **34**: 501-508.
- Turk, A. E. 1993. Articulatory phonetic clues to syllable affiliation: Gestural characteristics of bilabial stops, In P. A. Keating (ed.), *Papers in Laboratory Phonology III*, 107-135. Cambridge, UK: Cambridge UP.
- Walley, A. C., & Carrell, T. D. 1983. Onset spectra and formant transitions in adult's and child's perception of place of articulation, *J. Acoust. Soc. Am.*, **73**: 1011-1022.
- Walsh, M. A. 1991. Formant transition duration and amplitude rise time as cues to the stop-glide distinction, *Quat. J. Exptl. Psych., Hum. Exptl. Psych.*, **43A**: 603-620.
- Whalen, D. S., & Beddor, P. S. 1989. Connections between nasality and vowel duration and height: Elucidation of the Eastern Algonquian intrusive nasal, *Lg.*, **65**: 457-486.
- Williams, D. R. 1986. *Role of Dynamic Information in the Perception of Coarticulated Vowels*, Ph.D. dissertation, University of Connecticut.
- Wilson, C. 2001. Consonant cluster neutralization and targeted constraints, *Phonology*, **18**: 147-197.
- Wright, R. (1999). Tsou consonant clusters and auditory cue preservation, In E. Zeitoun & P. Li (eds.), *Selected Papers from the Eighth International Conference on Austronesian Linguistics*, v. 1, 227-312. Taipei: Academia Sinica, Institute of Linguistics.
- Wright, R. (2001). Perceptual cues in contrast maintenance, In E. Hume & K. Johnson (eds.) *The Role of Speech Perception in Phonology*, 251-277. San Diego: Academic Press.

Linguistics Department
University of Massachusetts
Amherst, MA 01003

jkingston@linguist.umass.edu