



PROJECT MUSE®

Phonetic Knowledge

John Kingston, Randy L. Diehl

Language, Volume 70, Number 3, September 1994, pp. 419-454 (Article)

Published by Linguistic Society of America

DOI: <https://doi.org/10.1353/lan.1994.0023>



➔ *For additional information about this article*

<https://muse.jhu.edu/article/452931/summary>

PHONETIC KNOWLEDGE

JOHN KINGSTON

*University of Massachusetts,
Amherst*

RANDY L. DIEHL

University of Texas, Austin

This paper argues that the phonetic interpretation of phonological representations may be controlled as well as automatic, because contextual variation in the realization of distinctive feature values is a flexible and adaptive response to variation in the demands on the production or perception of these values between contexts. The principal evidence presented in support of this argument is that the variation in the phonetic realization of speech sounds between contexts or languages involves reorganization of articulations into distinct phonetic categories. Extensive evidence of such reorganization in the realization of the feature [voice] is presented.*

'The tone and movement of my voice express and signify my meaning; it is for me to guide it to make myself understood . . . Speech belongs half to the speaker, half to the listener. The latter must prepare to receive it according to the motion it takes. As among tennis players, the receiver moves and makes ready according to the motion of the striker and the nature of the stroke.' Montaigne, *Of Experience* (1587–88 [1957]: 834).

INTRODUCTION

1.1. CONTROLLED VS. AUTOMATIC PHONETICS. This paper is about how in the act of speaking the phonetic component implements the strings provided by the phonological component. A common view of this act is captured in the metaphor of driving a car. The driver issues instructions to alter the car's motion through various control mechanisms. To the driver, the physical means by which these instructions are executed are irrelevant, so long as the car responds reliably to manipulation of the control mechanisms. However, the hidden workings of the car determine this response; e.g., the front wheels' freedom of movement and the car's length determine its turning radius. Also, environmental factors limit the car's response to the control mechanisms; e.g., the steering is constrained by the road's transverse angle. Though a successful driver knows about both the inner workings of the car and the environment in which it is driven, this is largely heuristic knowledge derived from empirical observations of how the car responds to manipulations of the control mechanisms.

Separating the driver and the control mechanisms from the working parts of the car and its operating environment is similar to separating the speaker's

* We are grateful to audiences at the University of Texas at Austin, Cornell University, the Ohio State University, William and Mary College, the Fourth Laboratory Phonology Conference, and the University of Massachusetts at Amherst for their insights and criticisms regarding the ideas presented here. The comments of Mary Beckman, Carol Krumhansl, Peter Ladefoged, Fred Landman, Anders Löfqvist, John McCarthy, Terry Nearey, Janet Pierrehumbert, Lisa Selkirk, Alice Turk, and three anonymous readers were very helpful as the paper passed through its protracted gestation. We also thank Christine Bartels for bringing the particularly apt quote from Montaigne's *Of experience* to our attention. Responsibility for the contents remains ours alone.

phonological intentions, as embodied by the paradigmatic and syntagmatic feature arrangements of which messages are composed, from their phonetic realization. In this view, the phonetic realization is determined by the vocal tract's anatomy and physiology as automatically as a properly designed and maintained car on a good road responds to its driver's manipulations of the control mechanisms.¹ In this paper, we argue that the phonetics is controlled rather than automatic.

Our use of 'automatic' and 'controlled' in describing phonetic implementation differs from that of certain other investigators in speech motor control and cognitive psychology. For example, Weismer & Cariski (1984) use 'control' in referring to how precisely and consistently a talker produces a specified phonetic outcome on demand, a much more stringent requirement than we will impose. An influential theory of cognitive processing (Schneider & Shiffrin 1977, Shiffrin & Schneider 1977) distinguishes 'automatic' and 'controlled' processes more generally: the former are unhindered by capacity limitations of short-term memory, do not require attention, run to completion mandatorily once initiated, require extensive training, are not easily modified, and are typically hidden from conscious perception; the latter, by contrast, are capacity-limited, attention-demanding, relatively easily learned and modified, and often accessible to conscious inspection. Most forms of phonetic implementation in the mature language user fall under the 'automatic' label in Schneider & Shiffrin's usage, whereas (we claim) they fall under the 'controlled' label in ours.

We focus in this paper on the covariation among articulations (and their acoustic consequences) that can be observed in any minimally contrasting pair of speech sounds. Our concern is whether some of the articulations automatically covary with others as a result of phonetic constraints, or if instead each of the covarying articulations is independently controlled. The cases we will discuss below can all be described as 'controlled' because we can show that language communities have selected the covariation from a larger set of phonetic possibilities.

We believe that phonetic implementation has acquired the appearance of an automatic process because it is so thoroughly overlearned. To acquire the fluency that makes phonetic implementation appear automatic, speakers must learn the appropriate control regimes for articulating each allophone of each phoneme.² Similarly, listeners must learn to recognize the patterns of acoustic

¹ This characterization misrepresents driving a car as much as it does speaking; in particular, the driver, like the speaker, uses knowledge of the car and its environment in much the same way that we will argue that the speaker uses knowledge of how the vocal tract and ear work. The traits that we will ascribe to speakers appear in fact to be general properties of action; whether it's driving or speaking, action is seldom mere execution.

² We employ the traditional terms 'phoneme', 'allophone', 'segment', and 'string' extensively in the discussion below as a convenient, largely theory-neutral shorthand. Though we don't believe our usage of these terms differs from the rest of the community's, it may be useful to be explicit about how we will use them here. PHONEME and PHONEMIC CONTRAST refer to any difference in the feature content or arrangement of an utterance's phonological representation which may convey a difference in semantic interpretation. This usage doesn't imply a distinct taxonomic phonemic

properties that characterize each allophone, as well as how the allophones correspond to phonemes. That the acquisition period is fairly lengthy and that various intermediate solutions are used during this period (Menn 1983, Nittrouer 1992) suggest that the fluent character of mature speaking and listening is a product of submerging highly controlled, but well practiced, behaviors below the level of conscious attention. Another reason speaking and listening occur below the level of conscious attention is that these acts are only a means to the end of conveying or understanding messages, and attention to message content will usually supersede attention to its phonetic medium.

In the following section we describe three points of view regarding the origins of the pervasive articulatory covariation observed in minimally contrasting speech sounds. These perspectives differ in how they interpret the metaphor of the 'physical speaking machine', to use Keating's apt phrase (1985, 1988).

1.2. MODELS OF THE PHYSICAL SPEAKING MACHINE. In a pair of important papers (1985, 1988), Keating characterizes the view of phonetic implementation proposed in Chomsky & Halle 1968 in terms of the metaphor of the physical speaking machine. We suggest that this metaphor has three interpretations, each embodied in a model of phonetic implementation. We will argue that only the third of these models adequately comprehends the facts of phonetic implementation.

1.2.1. A LITERAL AND INFLEXIBLE PHONETICS. The first is an essentially naive model of phonetic implementation that few phoneticians would espouse, in which the particular categories that occur in the surface phonological representation are translated into gradient vocal tract events in the same way every time they occur. That is, once the content of the surface phonological representation is known, the phonetic events in the vocal tract can be exhaustively and exactly predicted. Halle 1983 illustrates in detail how such a model would work, showing how vocalic distinctive feature values are mapped in just a single way onto the contraction of specific muscles. The naiveté of such models is demonstrated by the flexibility repeatedly observed in the translation of distinctive feature values into articulatory events (MacNeilage 1970, Ladefoged et al. 1972, Lindblom et al. 1979, Fowler & Turvey 1980, Kelso & Tuller 1983, Abbs et al. 1984).³

1.2.2. A FLEXIBLE BUT AUTOMATIC PHONETICS. More sophisticated models that acknowledge the flexibility with which distinctive feature values are

level of representation. ALLOPHONE refers to any phonetic variant of a distinctive feature specification or arrangement of such specifications that occurs in a particular context. SEGMENT is the term with the greatest potential to mislead; in our usage, this term refers to a collocation of articulations or acoustic properties which are organized so as to occur consistently together. This does not imply that a segment is a discrete event, all of whose articulations or acoustic properties are to be found within the same interval of time. Nor does it imply that the articulations or acoustic properties that participate in one collocation may not also participate in adjacent collocations. Finally, a STRING of segments is a set, possibly containing just one member, of such collocations.

³ Furthermore, the surface phonological representation is not translated literally into articulatory events (Ladefoged 1980, Wood 1982, Nearey 1978; cf. Fischer-Jørgensen 1985).

mapped onto articulations have generally been more appealing to phoneticians. The sophistication (and appeal) of these models rests in their recognition of three kinds of contingencies that influence this mapping.

The first contingency reflects the fact that successive speech sounds necessarily overlap with one another, sometimes quite extensively, rather than being discrete. The requirement of coarticulation (or coproduction) means that what articulations occur in implementing a particular distinctive feature value and how large they are will vary considerably with context. For example, speakers may not be able to retract the tongue body completely for a back vowel such as /u/ in the context of alveolar consonants, which demand a more forward position of the tongue body.⁴

The second contingency reflects the ways in which speakers accommodate the interactions between the distinctive feature values of the same speech sound. For example, the oral cavity must be expanded in a [+voice] stop (Westbury 1983) if oral air pressure during the stop is to rise enough to produce a noise burst when the stop is released without also rising so much as to shut down air flow through the glottis and thus voicing. The cavity expansion accommodates the aerodynamic incompatibility between [+voice] and [–sonorant], and is not required to implement either distinctive feature value alone.

A third contingency is physiological or mechanical dependences between one articulator and another, which bring about some articulations as automatic byproducts of other, intended articulations. For example, raising the tongue dorsum in the mouth to produce [+high] vowels is thought to change the tension of the vocal folds, causing them automatically to vibrate faster in [+high] vowels, either by pulling upward on the aryepiglottic folds (Ohala & Eukel 1987, Hombert 1978, Ewan 1979) or by advancing the hyoid bone and thereby tilting the thyroid cartilage forward (Honda 1983, Honda & Fujimura 1991, Sapir 1989).

In this more sophisticated model, phonetic implementation varies contingently rather than being invariant, as in the naive model above. Nonetheless, the variation is entirely predictable in that the surface phonological representation and the phonetic contingencies together completely determine the response of the phonetic component. The research strategy implicit in this model—one adopted by many phoneticians—assumes that the task of explaining recurrent articulatory covariation is merely a matter of identifying all the applicable contingencies.

As an illustration, consider Chen's 1970 observation that vowel duration differences before consonants contrasting for [voice] in English are larger than those observed in some other languages. This disparity in the size of these differences was apparently confirmed by Mack (1982), who showed that the differences were quite a bit smaller in French than in English. However, Laeufer (1992) has recently demonstrated that, once vowel durations are compared in

⁴ The articulatory and acoustic undershoot which results apparently presents no difficulty to listeners, since they expect it and can undo its effects so long as they hear the sounds that brought about the undershoot (Ohala 1981a).

prosodically matched contexts, the differences are equally small or large in the two languages (see also Davis & Summers 1989). This result suggests that Chen's and Mack's earlier work had simply failed to discover all the contingencies influencing the duration of vowels before consonants contrasting for [voice], rather than showing genuine crosslinguistic differences in the phonetic implementation of this contrast.

1.2.3. A CONTROLLED PHONETICS. Although this research strategy was ultimately successful in this case (but see Summers 1987 and Kluender et al. 1988 for remaining problems), there are other cases where it has failed so persistently that a different model of phonetic implementation might be worth considering, e.g. in explaining F_0 differences between vowels contrasting for height (for reviews see Silverman 1987, Sapir 1989, Fischer-Jørgensen 1990) or between vowels next to consonants contrasting for [voice] or other phonation features (for reviews see Hombert et al. 1979, Kingston 1985). (Lack of space prevents us from reviewing these failures in detail here, but see §3.3 below for some discussion.) The model we have in mind considers phonetic implementation to be governed by constraints that determine what a speaker (or listener) can do, but not what they must do; that is, the constraints limit phonetic behavior rather than predicting it.

A natural implication of such a model is that far more articulations are directly controlled by speakers than in either of the other two models considered. This sort of model therefore predicts more variability in the phonetic implementation of contrasts than the one that attributes all variability to contingencies between one articulation and another, so it would appear not to be restrictive enough to predict the observed crosslinguistic similarities in how articulations covary with one another.

Two considerations, one negative and the other positive, suggest that we should nonetheless adopt this apparently less restrictive model. The negative consideration is the fact that there actually is substantial variability between contexts, speakers, and languages in whether one articulation covaries with another in its size, and even in some cases its direction. There is in fact variation of a kind different from that predicted by a model in which all the variation is a product of contingencies among articulations. On the positive side, we will show below that a model in which speakers control most articulations achieves the necessary restrictiveness because it allows speakers to optimize their phonetic behavior, by both minimizing articulatory effort and maximizing perceptual distinctiveness (for similar views, see Lindblom 1983, 1990). Phonetic behavior is thus multiply constrained in such a model, and as much by the needs of the listener as by those of the speaker.

1.3. PHONETIC KNOWLEDGE. In arguing that phonetic implementation is not automatically determined by constraints reflecting articulatory contingencies, we will show that speakers and listeners employ extensive and subtle phonetic knowledge. Our claim that speakers adaptively control articulations is anticipated by Ohala (1981b), who has argued that because speakers recognize the phonetic consequences of such constraints on production or perception, they

may alter their behavior to compensate for or avoid these consequences. This heuristic knowledge 'feeds forward' on their speech acts, but shapes them differently than the 'feedback' from phonetic constraints, because the speaker chooses to produce a different sound and thus to avoid the consequences of the constraints.

These alternative choices may in time even alter the phonology of the language, as, for example, in the implosion of Common Indic geminate [+voice] stops in Sindhi (Nihilani 1974). In a geminate [+voice] stop, the closure probably lasts much too long for passive cavity expansion (Ohala & Riordan 1979) to keep intraoral air pressure (P_o) below subglottal air pressure (P_s). Imploding the stop recruits the oral cavity's considerable potential for active expansion (Westbury 1983) to ensure that the pressure difference and thus voicing is maintained through the long closure. The contrast in present-day Sindhi between implosives and voiced stops where geminate stops originally contrasted with single stops would come from uncoupling the large cavity expansion from the long closure, which is then dispensed with.⁵

Since no new articulations are added, this uncoupling exemplifies minimal exertion of phonetic control. We refer henceforth to adaptive changes in articulatory goals like this one, as well as to changes where new articulations are employed, as 'reorganizations' of articulations. The reorganizations can furthermore be qualitatively as well as quantitatively different from one another, so they are not uniquely predicted even from motor equivalence constraints which allow flexibility in the use and size of the articulations that contribute to an articulatory goal (cf. Fowler et al. 1980).

The phonetic knowledge that underlies these reorganizations is intimately connected to the phonology of the language, because it is by means of this knowledge that phonological strings are transformed into articulations or are recognized in the acoustic signal; that is, it is knowledge about phonological representations as well as phonetic constraints.

But despite the fact that a good bit of what a speaker or listener knows phonetically can be traced back either to the language's phonology or to phonetic constraints, together these two still do not uniquely predict the speaker's phonetic behavior. Variation might be automatic if each variant were the sole realization given either the context in which it occurs or the language's phonology, because such variation could then be represented in the phonology or, if it was in the phonetics, it might be predicted entirely from phonetic constraints.

But elevating noncontrastive contextual or crosslinguistic variation out of the phonetics into the phonology would divorce the variation from its phonetic motivations, allow variation that never occurs, and reduce an explanation of the variation to mere description. Furthermore, treating this variation in the

⁵ Confronted with the same aerodynamic difficulty as is presented by Sindhi's [+voice] geminates, speakers of other languages have chosen to succumb to the difficulty and devoice them, while yet others simply shorten them (Jaeger 1978, Ohala & Riordan 1979). Of course, many languages keep their [+voice] geminates, presumably by making maximal use of the oral cavity's capacity to expand actively.

phonology would eliminate a distinction between the phonological processes which change distinctive feature values and phonetic processes which implement them. However, even though this variation is phonetic, a unique outcome is not predicted automatically from the phonetic constraints, which only limit the range of alternatives from which speakers may choose. Because in any particular instance this choice is not entirely determined by the joint effects of phonological representations and phonetic constraints, the component in which this knowledge is represented cannot be automatic but must instead be a controlled mechanism for implementing phonological contrasts.

Like us, Keating (1990) treats allophonic variation as controlled phonetic differences. Where our views converge with Keating's (and where both differ from her earlier proposals [1985, 1988]) is in not putting any phonic process or attribute of an utterance which is specific to particular languages rather than universal into their individual phonologies. Contrastiveness rather than language-specificity should be the criterion for assigning a phonic attribute of an utterance to the phonology.

With this division of labor between phonetics and phonology, the phonology regains a level of abstractness appropriate to stating such rules as voicing assimilation, which is indifferent to contextually (or segmentally) determined variation in the realization of the feature [voice]. Also, a distinction is regained between phonetic attributes which have been 'phonologized' and those which are simply controlled (see Kingston & Solnit 1988, 1989). Finally, this division means that only phonological features, and not also phonetic properties, may be distinctive or redundant (cf. Stevens et al. 1986). Phonetic properties may differ instead in how reliably each occurs in the speaker's output and how salient the listener finds each as a cue to a contrast.

The remainder of this paper is structured as follows. In §2 we will show that contextual variability in the phonetic realization of [+voice] stops reflects quantifiable differences between contexts in the effort required to make the vocal folds vibrate. We argue that speakers of some languages respond to those differences by synchronically reorganizing the articulations they produce so as to minimize articulatory effort. Speakers' implicit understanding of differences in the energetics of voicing between contexts reflects what we'll call 'speaker-oriented' phonetic knowledge.

In §3 we consider the fact that F_0 is depressed next to [+voice] stops regardless of what other articulations occur in their production in different contexts or languages. The consistent depression of F_0 might suggest that the articulation of [+voice] stops is fundamentally the same in all contexts and languages, rather than being reorganized to save effort. We argue instead that, because no other articulation is likely to produce the F_0 depression as an automatic by-product, the depression must itself be a product of an independently controlled articulation, whose purpose is to enhance the [voice] contrast.

Finally, §4 presents an array of evidence suggesting that articulations may be covaried by speakers because their acoustic consequences interact in such a way as to enhance minimal contrasts. In a model where the covariation is entirely a product of articulatory contingencies rather than independently con-

trolled articulations, the combinations of acoustic consequences produced by the covariation are instead quite arbitrary, and any perceptual interactions among them are fortuitous. The implicit understanding that leads speakers to covary articulations whose acoustic consequences enhance one another perceptually is 'listener-oriented' phonetic knowledge.

Before proceeding, we should note that we have chosen to discuss [voice] because it has been extensively studied and because its phonetic implementation clearly exhibits the kind of control that, we argue, speakers exercise. We omit consideration of other features not because their phonetic implementation is less controlled (see Kingston 1991 for evidence of such control in vowel height contrasts, for instance), but because we have no space to discuss them.

SPEAKER-ORIENTED PHONETIC KNOWLEDGE

2.1. VARIATION IN THE REALIZATION OF THE FEATURE [voice]. Sounds which contrast for the phonological feature [voice]⁶ are not realized in the same way in all contexts. For stops in English, at least the realizations laid out in Table 1 can be distinguished⁷ (see e.g. Peterson & Lehiste 1960, Lisker & Abramson 1964, Stevens & Klatt 1974, Lisker 1957, 1986 [1978], Hombert 1978, Docherty 1989).

Furthermore, the [voice] contrast collapses after /s/ to a voiceless unaspirated stop and through flapping in alveolar stops before an unstressed syllable (cf. Fox & Terbeek 1977). This variation is great enough to suggest that these phonemes (or distinctive feature values) may have no essential phonetic properties, since an allophone of a phoneme in one context may share few or no transcribable properties with allophones of that phoneme elsewhere.

2.2. IS IT THE SAME CONTRAST? The variation laid out in Table 1 also raises the question of how it's decided that two languages contrast two sounds for the same distinctive feature, since stops differ crosslinguistically along many of the same dimensions in which they vary with context (see also Keating et al. 1983 and Keating 1984 for relevant discussion).

For example, among the Germanic languages, contrasting stops differ in glottal-oral timing (and consequently VOT [voice onset time]) in at least three

⁶ The feature specifications [+voice] and [-voice] refer here only to the phonological feature [voice]. The terms 'voiced', 'voiceless', 'voicing', and 'vocal fold vibration' (these last two interchangeably) refer here to whether the vocal folds actually vibrate, i.e. to a phonetic property of the utterance.

⁷ While [-voice] stops are realized in much the same way word-initially as utterance-initially, the definition of initial position is problematic for [+voice] stops, since their word-initial allophones, when produced by some speakers or in some prosodic contexts, resemble intervocalic allophones in having closure voicing when the preceding word ends in a vowel (Lisker & Baer 1984). For other speakers or contexts, however, even after words ending in a vowel, word-initial [+voice] stops resemble utterance-initial ones. Caisse 1982 and Docherty 1989 present evidence that both American and British English speakers' word-initial as well as utterance-initial [+voice] stops are most often voiceless unaspirated, even when the preceding word ends in a vowel. Therefore, in the discussion below, data from word-initial stops will be treated together with data from utterance-initial ones when their phonetic realizations are substantially similar.

	[+ voice]	[− voice]
UTTERANCE-INITIAL OR PRETONIC;	short lag VOT F ₁ lower F ₀ lower weaker burst	long lag VOT F ₁ higher F ₀ higher stronger burst
INTERVOCALIC OR POSTTONIC	closure voicing shorter closure longer preceding vowel F ₁ lower F ₀ lower	no closure voicing longer closure shorter preceding vowel F ₁ higher F ₀ higher
UTTERANCE-FINAL AND POSTVOCALIC	longer preceding vowel closure voicing possible shorter closure F ₁ lower	shorter preceding vowel no closure voicing longer closure F ₁ higher

TABLE 1. Partial descriptions of phonetic variants of English stops contrasting for [voice]; the differences in F₁ and F₀ occur at the edges of adjacent vowels.

different ways. First, English, Swedish, and German contrast voiceless aspirated stops with unaspirated or occasionally prevoiced stops initially and voiceless unaspirated stops with prevoiced ones intervocalically and posttonically. Second, Icelandic contrasts voiceless aspirated stops with an unaspirated series that is never voiced during the closure (Petursson 1976). Third, Dutch contrasts voiceless unaspirated stops with regularly prevoiced stops (Slis & Cohen 1969a,b, Slis 1970, Collier et al. 1979). Can all of these languages be said to distinguish stops for the feature [voice], or is this true only of the Dutch stops, while in Icelandic the distinctive feature is [spread glottis] and in English, German, and Swedish yet a third laryngeal feature? Despite their differences, there are both phonetic and phonological reasons to favor [voice] in all five languages.

First, in all five, voicing begins earlier relative to the stop release in one series than another (because the glottis is closed⁸ earlier). This suggests that [+voice] can be applied crosslinguistically to the stop series in which voicing begins earlier and [−voice] to the one where it begins later (see Lisker & Abramson 1964, 1971, Keating 1984, Keating et al. 1983).

Second, as shown in §3 below, F₀ is consistently depressed in vowels next to [+voice] stops, regardless of whether a language or context employs a prevoiced or short lag stop as the realization of this phonation type. We show there that F₀ varies only with a [voice] contrast and not simply with the presence vs. absence of phonetic voicing. In §4 we argue that F₀ depression next to a

⁸ In this discussion, the glottis is said to be 'closed' when the vocal folds are brought into medial contact by rocking the arytenoids together through interarytenoid contraction. Medial contact is sufficient to impede air flow through the glottis enough to cause subglottal air pressure (P_g) to rise. This rise in P_g eventually blows the folds apart, allowing air to flow up between them (U_g⁺). This flow creates the pressure drop in the glottis itself known as the Bernoulli effect, which, combined with sustained interarytenoid contraction, pulls the folds back into medial contact. This state of the folds thus closes the glottis only periodically, unlike in a glottal stop, in which additional medial compression of the folds brought about by vocalis contraction yields a sustained, airtight glottal closure.

[+ voice] stop enhances the perceptual effect of earlier onset of voicing, because both contribute to the perception of low-frequency energy in the stop's vicinity. That languages such as the five Germanic languages just discussed should enhance a variety of differences in relative glottal-oral timing with similar F_0 differences is further evidence that the contrast involves the same distinctive feature.

Finally, despite the readily observed differences in how a putative [voice] contrast is realized in Dutch and English stops, stops (and other obstruents) assimilate in laryngeal articulation in clusters in both languages.

These arguments not only support the contention that the distinctive feature is the same—i.e. [voice]—in all five languages, but also show that it's possible and even essential to argue for a specific if relatively abstract phonetic content for distinctive features (cf. Coleman 1992).

2.3. THE FEATURE [voice] IN INITIAL AND INTERVOCALIC CONTEXTS. Explaining the difference between initial and intervocalic contexts rests on the greater difficulty of initiating than continuing vocal fold vibration (Lindqvist 1972, Westbury & Keating 1980). Aside from the necessity for the vocal folds to be sufficiently close together, their vibration depends on an outward flow of air between them (U_g^+), which comes from a minimum difference between subglottal air pressure (P_s) and intraoral air pressure (P_o), of 1–2 cm H_2O (Lindqvist 1972, Müller & Brown 1980). The P_o buildup behind a stop closure quickly reduces this P_s-P_o difference. If initiating voicing during a stop closure requires a greater P_s-P_o than maintaining voicing into a stop from a preceding vowel (Lindqvist 1972, Westbury & Keating 1980), voicing is less likely to begin in initial stops before the stop is released, but can more easily continue into and perhaps all the way through an intervocalic stop.⁹

In fact, intervocalically, it is turning voicing off that's more difficult, since the vocal folds won't stop vibrating until the glottal aperture (A_g) is substantially larger than what is needed to initiate vibration (Lindqvist 1972, Hirose et al. 1985). This hysteresis undoubtedly arises in part from the larger P_s-P_o needed to initiate than to maintain voicing.¹⁰ In the transition of the folds from a non-vibrating to a vibrating state, A_g must get quite small before the resistance to transglottal airflow will rise enough to elevate P_s above P_o , but so long as the folds are already vibrating, they apparently can continue to do so until A_g gets rather large.

⁹ A reviewer observes that the pressure drop required to initiate voicing may only be slightly larger—about 2 cm H_2O —than that needed simply to maintain voicing. But if the drop needs only to be increased modestly to initiate voicing, perhaps through active oral cavity expansion, then voicing during the closure should be more common in initial stops than it is, especially given the fact that for many speakers the folds are (appropriately) adducted in initial [+ voice] stops (Flege 1982).

¹⁰ A reviewer notes that because the vocal tract walls are compliant, the oral cavity will expand passively once the oral closure is complete, slowing the rate of increase in P_o . This effect acts synergistically with the smaller P_s-P_o difference needed to maintain voicing from a preceding vowel to allow it to persist well into a following closure.

In summary, the requirements on A_g as well as U_g^+ which must be met to initiate voicing are more stringent than those for simply maintaining it. Note that we do not appeal here to some general, vague notion of greater difficulty or effort, but instead to quantifiable differences in (i) the size of the U_g^+ needed to initiate vs. maintain voicing and (ii) the A_g needed to extinguish vs. initiate it.

In an automatic phonetics, the glottis would be closed during the closure in both initial and intervocalic [+voice] stops, and voicing would be much more common during intervocalic than during initial closures entirely because of the lesser U_g^+ needed to maintain than to initiate voicing. Initial [+voice] allophones would most often be voiceless unaspirated, as an accidental consequence of not expanding the oral cavity and reducing P_o enough to elevate U_g^+ above the voicing threshold. By contrast, phonetic control would allow the speaker, in an initial [+voice] stop closure, to keep the glottis open until the stop release, when U_g^+ becomes high enough to initiate voicing without help from other articulations. The attempt to produce voicing during closure is deliberately abandoned. In intervocalic [+voice] stops, both an automatic and a controlled phonetics predict that the glottis will be kept closed from the preceding vowel into the closure, since maintenance of voicing requires only a small U_g^+ .

Unlike a controlled phonetics, an automatic phonetics also predicts that [-voice] allophones shouldn't vary between initial and intervocalic positions. A voiceless aspirated allophone may be demanded for the [-voice] stop initially to avoid confusion with the frequent voiceless unaspirated allophones of the [+voice] stops,¹¹ but in an automatic phonetics there's no reason why it shouldn't occur intervocalically, too. In a controlled phonetics, however, the reliable closure voicing of intervocalic [+voice] stops allows the aspiration required of initial [-voice] stops to be dispensed with. Since the [-voice] stops must be just sufficiently less voiced than the [+voice] stops in each context, the timing of the laryngeal articulation relative to the oral one used for [+voice] stops initially may be used for [-voice] stops intervocalically.

The question, then, is whether the articulations of either [+voice] or [-voice] stops are organized differently in intervocalic than initial position, along the lines predicted by a controlled phonetics. At first glance the answer appears to be 'no,' since the majority of Swedish and English speakers close the glottis substantially before the stop release in initial [+voice] stops (Lindqvist 1972, Flege 1982). Flege (1982) in fact found that only two of the ten American English speakers he examined left the glottis open until the stop release in utterance-initial [+voice] stops, while the remaining eight always closed the glottis completely in initial [+voice] stops before voicing began. That there are even two speakers who didn't close the glottis until the stop

¹¹ A more general principle may be at work here. The realization of the [-voice] member of the contrast may be prohibited from having an earlier onset of voicing than the [+voice] stop it contrasts with, since this would contradict the phonological specification of the contrast (see Pierrehumbert 1980, Kohler 1984).

release is enough to support the claim that speakers can retime glottal closure in response to the greater aerodynamic difficulty of initiating voicing.

But what of the majority, who didn't? All eight actually closed the glottis long before the oral closure was made (and similar data are reported for Swedish speakers by Lindqvist 1972). This glottal closure may, therefore, be part of assuming a general speech posture, rather than being intended specifically for the initial [+voice] stop. This interpretation is supported by Lindqvist's observation of glottal closure before speech began even in some tokens where the initial segment was [-voice], which would require opening the glottis after closing it.

Voicing during the stop closure itself may therefore require additional mechanisms beyond closing the glottis, particularly oral cavity expansion, to ensure adequate U_g^+ . That the cavity is not always expanded is shown by the frequent voiceless unaspirated allophones produced by the three of Flege's eight speakers who always closed the glottis long before the stop release. Voiceless unaspirated allophones were also typical in initial [+voice] stops produced by Lindqvist's Swedish speakers, again despite the glottal closure.

Caisse's 1982 study of five American English speakers and Docherty's 1989 study of five British English speakers both exaggerate these trends, since in both studies voiceless unaspirated allophones were markedly more common than prevoiced ones for word-initial [+voice] stops. Docherty's data are especially striking here, because voiceless unaspirated allophones were still the clear favorite even when the preceding word ended in a vowel. This overwhelming preponderance of voiceless unaspirated allophones raises the possibility that some of the speakers in these studies left the glottis open until the stop release.

Intervocally, by contrast, the glottis is reliably closed during the closure of intervocalic [+voice] stops in English, Swedish, and probably German,¹² and closure voicing is much more likely there in all three languages (Lindqvist 1972, Löfqvist 1975a,b, 1976, Löfqvist & Yoshioka 1980, 1981a, Caisse 1982).

These data show that speakers choose between two active articulations in producing initial [+voice] stops. Some delay glottal closure until the stop release and avoid the difficulty of initiating voicing. Others, who close the glottis, may also expand the oral cavity to overcome this difficulty, though they occasionally or frequently fail to produce this expansion.

There is nonetheless reason to doubt whether the contextual variation between initial and intervocalic contexts in the articulation of [-voice] stops requires reorganization. Browman & Goldstein (1990, 1992; see also Fowler 1990, Pierrehumbert 1990, Sproat & Fujimura 1993) have argued recently that, rather than producing the distinct phonetic categories implied by reorganiza-

¹² Klaus Kohler (personal communication, 1988) describes the contrast in German as one of aspiration both initially and intervocally, but, as in English, the [-voice] stops have less aspiration intervocally. German [+voice] stops may or may not have voicing during the closure intervocally. The complete absence of any initial prevoiced allophones in the data from Caisse's (1982) two German speakers suggests that they, too, may delay closing the glottis until the stop release, as did the two of Flege's English speakers who produced only voiceless unaspirated allophones of [+voice] stops initially.

tion, allophonic variation is instead continuous variation in the size and timing of the gestures that realize a distinctive feature value, or in their overlap with the gestures of adjacent segments.¹³

In describing the difference in the realization of English [–voice] stops between initial and intervocalic position, Browman & Goldstein (1992) appeal to data from two speakers of American English in Cooper 1991a as showing that the glottis is always opened in [–voice] stops in English—intervocalically as well as initially and before unstressed as well as before stressed vowels. Glottal opening is simply smaller intervocalically than initially and before unstressed than before stressed vowels, and this smaller opening leads to shorter voicing lags (VOTs) and thus less aspiration. Furthermore, Browman & Goldstein suggest that the effects of position and stress on the size of the glottal opening are additive.¹⁴ These results would follow from continuous variation in the size and timing of the glottal opening rather than from categorical effects of word position and following stress.

However, the condensed account of Cooper's data in his 1991b paper shows that intervocalic stops have much longer VOTs before a stressed vowel than before an unstressed vowel, while initial stops have relatively long VOTs regardless of the stress of the following vowel. This result is surprising given that initial position as well as following stress increases the glottal opening's size and duration. Since the oral closure's duration is also increased by both of these effects, we might expect VOTs to remain more or less the same across manipulations of word position and following stress. Apparently, speakers have adjusted some other aspect of the glottal or oral articulations to produce the differences in how VOTs are affected by stress intervocalically and initially.

Cooper (personal communication, 1994) has provided evidence from the relative timing of glottal and oral articulations that suggests what this adjustment might be. First, in intervocalic stops the interval from the onset of glottal opening to the onset of glottal closing covaries much more with the duration of the oral closure when the following vowel is stressed than when it is unstressed. Second, in initial stops, following stress does not affect the extent of this covariation when the stop is labial, and it affects the covariation to a lesser extent than intervocalically when the stop is lingual. In intervocalic stops, speakers thus delay the onset of glottal closing to a greater extent before stressed vowels to produce the longer VOTs observed there, while initially, the onset of glottal closing is delayed (nearly) as much before unstressed as before stressed vowels, and VOTs are equally long before both kinds of vowels.

The extent to which peak glottal opening (= the onset of glottal closing) covaries with oral closure duration depends on stress in Swedish voiceless stops

¹³ We agree with Browman & Goldstein that allophonic processes should not be described as manipulations of distinctive features, which should be reserved for representing phonological contrasts and the statement of phonological rules (see above), but we do not believe this precludes the kind of categorical differences between allophones implied by our notion of reorganization.

¹⁴ In Table 1, what we've been referring to as a contrast between initial and intervocalic position confounds position and stress effects. Their effects are unconfounded in the present discussion.

in much the same way that it does in English, covarying more before stressed vowels than before unstressed vowels (Löfqvist 1980, Löfqvist & Yoshioka 1981a, 1984; cf. Löfqvist 1986, 1990). This difference, when combined with smaller peak openings, produces variation in VOTs with stress and word position in Swedish much like that observed in English.

The relative glottal-oral timing in these unaspirated [–voice] stops in English and Swedish is in fact similar to that of the unaspirated stops of other languages, e.g. Icelandic (Petursson 1976, Löfqvist & Yoshioka 1981b) and Hindi (Kagaya & Hirose 1975, Benguerel & Bhatia 1980; cf. Dixit 1989). Our account predicts correctly that a categorical difference in the relative timing of glottal and oral articulations which is used consistently in these other languages could also be found between contextual allophonic variants, while Browman & Goldstein's does not. (Another case where contextual variation in glottal-oral timing in [–voice] stops turns out to be categorical is presented in Tuller & Kelso 1990; but see Munhall & Löfqvist 1992 for a potentially more problematic case.)

In the next section we turn our attention to a phonetic property of stops contrasting for [voice] that does not appear to vary with context or language in the way VOT does; this is the lower F_0 in vowels next to [+voice] stops than in vowels next to [–voice] stops.

THE CONTROL OF F_0 NEXT TO STOPS CONTRASTING FOR [voice]

3.1. F_0 DIFFERENCES ARE INDEPENDENT OF GLOTTAL APERTURE. F_0 is uniformly depressed next to [+voice] stops, regardless of how the [voice] contrast is otherwise realized. This uniformity might suggest that [+voice] stops have the same glottal articulation in different contexts or languages after all, rather than being produced with different organizations of articulations. But closer scrutiny indicates that this interpretation is mistaken, and furthermore that the F_0 differences are a product of articulations that are controlled independently of the timing and size of the glottal articulation.

Regarding the effects of context, F_0 is lower in vowels next to initial as well as intervocalic [+voice] stops, regardless of whether voicing is present during the closure (Caisse 1982). Data from American English speakers similar to the data obtained by Caisse can be seen in Figure 1a–c (Kingston 1985 and unpublished data).

The two speakers in Figure 1a–b uniformly produced voiceless unaspirated realizations of the [+voice] stop; nonetheless, as these figures show, the following F_0 contour is uniformly lower following this realization of a [+voice] stop than after the aspirated allophone of the [–voice] stop for both speakers (the differences between the means are small but there is also very little overlap between the ranges).

The speaker in Figure 1c also produced a substantial majority of voiceless unaspirated allophones for [+voice] stops, and again F_0 is uniformly lower in the following vowel than after the voiceless aspirated allophones of her [–voice] stops.

F_0 is also reliably lower after the prevoiced allophones of such stops produced consistently by the two speakers in Figure 2 (Kingston 1989), as after [+voice] stops in languages such as Spanish (Caisse 1982), which always have voicing

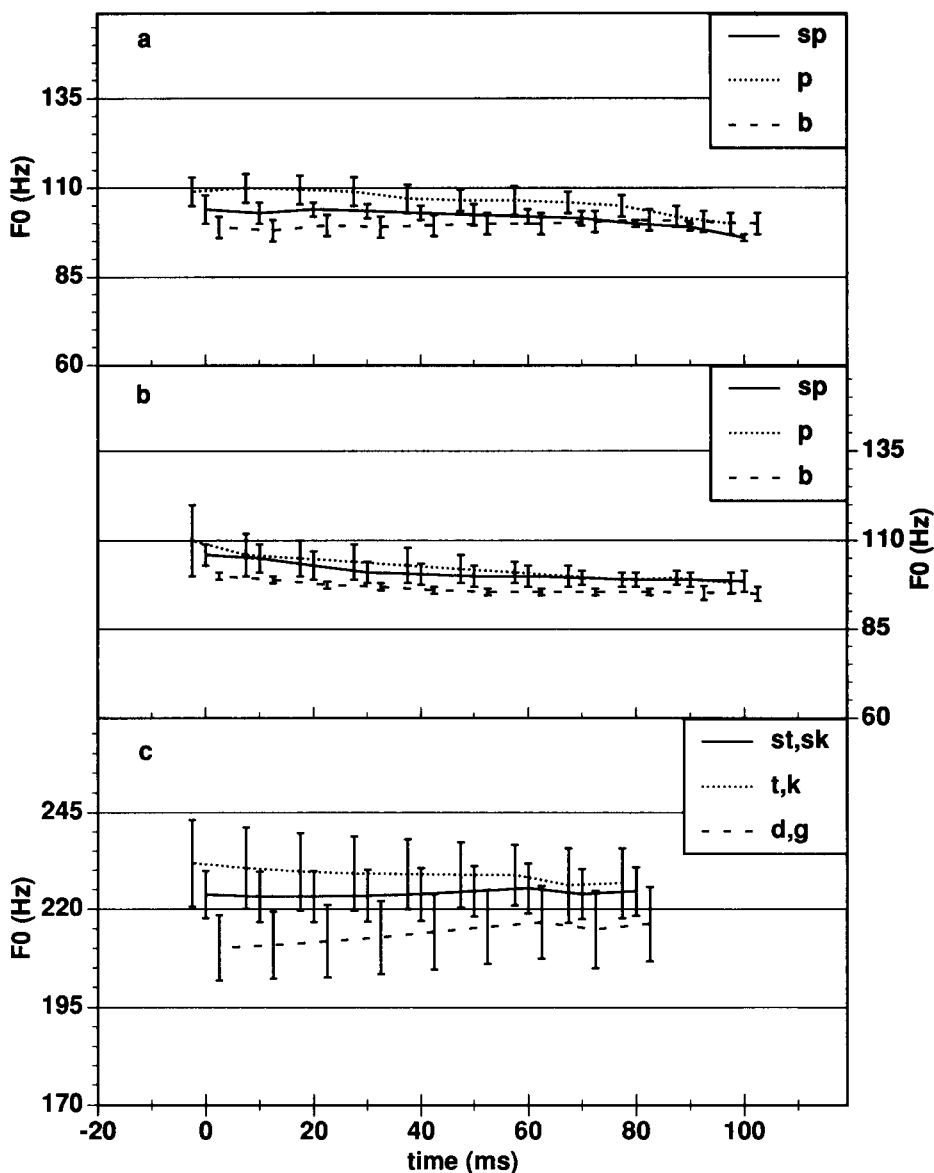


FIGURE 1. a–b: Mean F_0 contours following word-initial [(s)p, p^h, p] the phonetic realizations of /s)p, p, b/, respectively (the phonological representations are used in the figure legend), before [A] produced by two adult male speakers of American English. Error bars indicate one standard deviation (the curves are staggered slightly with respect to the horizontal axis so that the size of the error bars can be clearly seen in each case). $n = 6-7$ repetitions of each utterance type. c: Mean F_0 contours following word-initial [(s)t, (s)k, t^h, k^h, b, d], the phonetic realizations of /s)t, (s)k, t, k, b, d/, respectively, where the vowel was one of [i, u, A], produced by an adult female speaker of American English (error bars are 1 standard deviation with respect to this average). $n = 20$ repetitions for each consonant-vowel combination. Solid line = [(s)p, (s)t, (s)k], dotted = [p^h, t^h, k^h], dashed = [p] for a–b and [d, g] for c. F_0 values were obtained by measuring the duration of periods beginning with the first immediately following the consonant, and continuing well into the vowel. F_0 values were then interpolated at 10 ms intervals, and the resulting contours averaged for these displays. Time = 0 in these plots corresponds to the onset of voicing, i.e. the first period of the vowel.

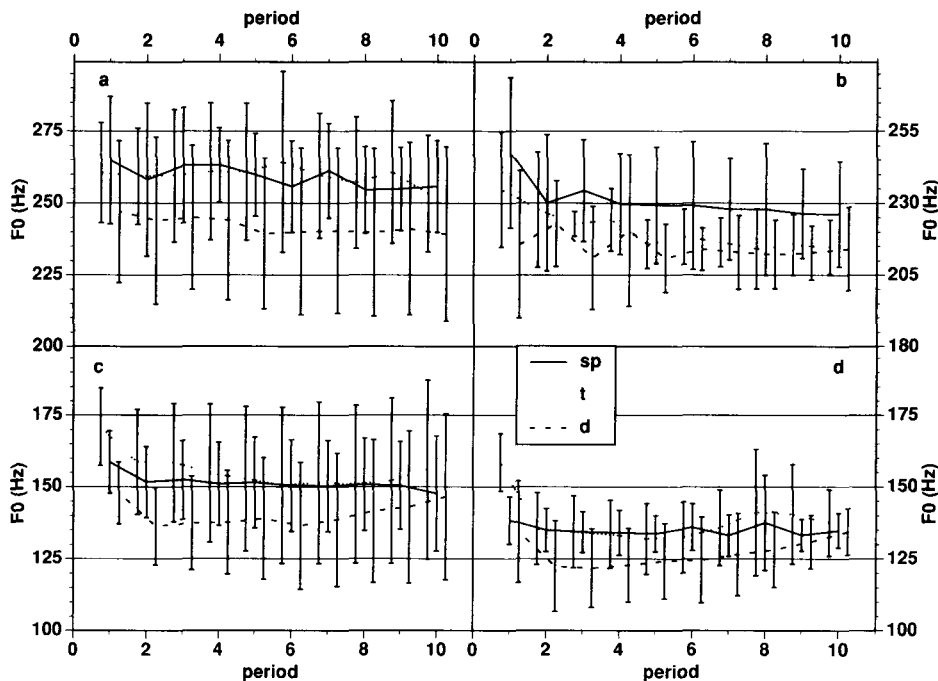


FIGURE 2. Mean F_0 contours (error bars = ± 1 standard deviation) for the first 10 periods of vowels following word-initial [t^h , (s)p, d] before primary-stressed and unstressed vowels, produced by an adult female speaker (a–b) and an adult male speaker (c–d) of American English. The words were either in focus (a,c) or out of focus (b,d). Solid line = [(s)p], dotted = [t^h], dashed = [d]. $n = 10$ –12 for each contour. Again, F_0 was obtained by measuring period durations.

during closure (Lisker & Abramson 1964). The invariant lowering of F_0 implies that some laryngeal articulation in English [+voice] stops is the same across initial and intervocalic contexts.

What is this articulation, and could it depress F_0 automatically in adjacent vowels? At first blush, the most likely candidate is glottal closure, but the available data, presented in Table 2, show that glottal closure alone is sufficient

	[voice]	[spread glottis]
English, German, Swedish, Korean	$p \sim b < p^{(h)}$	
Danish	$p \sim b = p^{(h)}$ $p \sim b < p^{(h)}$	
Spanish, French, Italian, Portuguese, Japanese	$b < p$	
Hindi	$b < p$	$p > p^h$
Mandarin, Fukieneese, Cantonese		$p > p^h$
Thai	$b < p$	$p > p^h$ $p < p^h$

TABLE 2. The direction of F_0 differences next to prevoiced [b], voiceless unaspirated [p], and voiceless aspirated [p^h] stops. Two entries for a language indicate that there are differences between speakers.

neither to produce voicing nor to lower F_0 , and that F_0 lowering may instead come from a deliberate slackening of the vocal folds. The most troublesome fact is that a small glottal aperture does not always lower F_0 , nor is F_0 always lower next to a stop with a smaller glottal aperture than next to one with a larger aperture.¹⁵

As suggested by the data in Table 2, it might be more reasonable to analyze the /p/:/p^h/ contrast as one of [spread glottis], with unaspirated stops [– spread glottis] in contrast to the aspirated stops [+ spread glottis], rather than [voice]. When a voiceless unaspirated stop represents the [– spread glottis] category, it usually elevates F_0 relative to the [+ spread glottis] stop, but it may also depress F_0 . These data thus show that F_0 is not necessarily lower in vowels next to the stop with the smaller glottal aperture. The data in Table 2 further suggest that voiceless unaspirated stops which realize a member of a [voice] contrast may both elevate and depress F_0 (see also Hombert & Ladefoged 1976). This result is paradoxical only if one ignores the phonological specification of this phone: when it represents the [– voice] category, as in Hindi, Thai, Spanish, French, Portuguese, Italian, and Japanese, then F_0 is elevated; but when it represents the [+ voice] category, as in English, German, Swedish, Danish, and Korean, then F_0 is depressed.¹⁶

3.2. F_0 DIFFERENCES DEPEND ON [voice]. In this section we present two further pieces of evidence which suggest that F_0 differences are predictable only from

¹⁵ The sources for Table 2 are: English and German: F_0 : Caisse 1982; Swedish: F_0 : Lofqvist 1975a, glottal articulations: Lofqvist 1972, Lofqvist 1975b; Danish: F_0 : Reinholt Petersen 1978, 1983, Fischer-Jørgensen 1968, Jeel 1975, glottal articulations: Frøkjær-Jensen et al. 1971, Fukui & Hirose 1983, Hutters 1985; Korean: F_0 : Han & Weitzman 1970, Hardcastle 1973, glottal articulations: Kim 1965, 1970, Kagaya 1971, 1974, Hirose et al. 1981; Hindi: F_0 : Dixit & MacNeilage 1980, glottal articulations: Kagaya & Hirose 1975, Benguerel & Bhatia 1980, Dixit 1989; Mandarin: F_0 and glottal articulations: Iwata & Hirose 1976; Fukiene: F_0 and glottal articulations: Iwata et al. 1979; Cantonese: F_0 and glottal articulations: Iwata et al. 1981; Thai: F_0 : Gandour 1974, Erickson 1975, Ewan 1976, glottal articulations: Rischel & Thavisak 1984; Spanish, French, Italian, and Portuguese: F_0 : Caisse 1982; and Japanese: F_0 : Kawasaki 1983, glottal articulations: Sawashima et al. 1975, Hirose & Ushijima 1978, Sawashima & Hirose 1983, Yoshioka et al. 1982.

In initial position in Korean, the difference is between stops with slight aspiration and stops with heavy aspiration, i.e. moderate vs. long lag VOTs. The former are the initial realization of Korean's so-called 'lax' stops in this context. The third series of stops in Korean, the so-called 'tense' stops, have the glottis closed at the stop release (Kagaya 1974), but F_0 is still quite high in following vowels (Han & Weitzman 1970, Hardcastle 1973). This, too, is evidence that glottal closure doesn't invariably lower F_0 , though in this case the 'unexpected' F_0 elevation can be attributed to the elevated fold tension produced by vocalis contraction (Hirose et al. 1981). Since F_0 is elevated next to other kinds of stops in which the vocalis is not active, this mechanism is specific to Korean.

In the Chinese languages Mandarin, Fukiene, and Cantonese, the high F_0 after voiceless unaspirated stops is especially noteworthy given the very small glottal opening in these stops relative to that observed in the aspirated stops of these languages.

¹⁶ The contrast between unaspirated (or weakly aspirated) and aspirated stops in both Korean and Danish is analyzed here as a [voice] rather than a [spread glottis] contrast, because in both languages the unaspirated stops are voiced intervocally, as the English [+ voice] stops are. Danish actually collapses the contrast there to a voiced stop. The Korean lax stops can also be voiced at the beginning of words within phonological phrases (Silva 1991, 1992), and voicing is inhibited within words after a devoiced vowel (Jun 1992).

the phonological specification for [voice] and not from other phonetic attributes of stops.

3.2.1. VARIABLE EFFECTS ON F_0 FOLLOWING ENGLISH [s]+STOP CLUSTERS. That it is contrast for [voice] in English consonants that determines F_0 in adjacent vowels, rather than their other phonetic properties, is indicated by the quite variable effects on F_0 of the voiceless unaspirated stops following [s]. This variability is expected since after [s] the [voice] contrast has been neutralized and can no longer exert any control over F_0 .

In some reports (Caisse 1982, Ohde 1984), F_0 is as high following the unaspirated stop after [s] as it is following the aspirated allophone of a [-voice] stop occurring alone, but this is by no means always true. For the three speakers in Fig. 1a–c, the mean F_0 contour following [(s)p], in which the [voice] contrast is neutralized, lies between the contours following the stops that contrast for [voice]. One of the two speakers in Fig. 2 (Fig. 2c–d) resembles the speakers in Fig. 1 in that the F_0 contour following the neutralized stop lies between those for the stops contrasting for [voice], while the other (Fig. 2a–b) resembles speakers described in Caisse 1982 and Ohde 1984.

3.2.2. F_0 DOESN'T DEPEND ON PHONETIC DIFFERENCES IN VOICING IN TAMIL STOPS. Finally, F_0 differences are not always observed next to stops that differ in whether there is voicing during the closure, as in Tamil, where nongeminate stops are predictably voiced but geminate stops are voiceless. Because voicing is predictable from length in Tamil stops, the data presented below from this language are a strong counterexample to the claim that F_0 differences in adjacent vowels are an automatic consequence of the stops' other glottal articulations. Instead, these data suggest that F_0 will only vary with the presence of voicing in stops that contrast for [voice] (see also Kohler 1982, 1984, 1985).

Before the Tamil data can be taken as a valid counterexample, however, it must be demonstrated that consonant [length] rather than [voice] is what is contrastive for Tamil stops. Lisker (1958) analyzes the Tamil contrast as one of [voice] rather than of [length] because, for a rather small number of tokens from a single speaker, he found a consistent difference in voicing, but large variation in duration, from very short to quite long, for the geminate stops (see also Balasubramanian 1980). Kingston's 1986 data from three Tamil speakers (T_1 – T_3) on the durations of intervocalic stops and sonorants, in Table 3, differ.

The data in Table 3 are drawn from three speakers and include many more tokens than in Lisker's study; they show that relative differences in closure

SPEAKER	SEGMENTS	GEMINATE STOP	SINGLE STOP	GEMINATE/SINGLE SONORANT
T_1	tt:d:ɳɳ	140 (10)	57 (11)	182 (19)
T_2	tt:d:ɳɳ	139 (10)	37 (8)	139 (12)
T_3	tt:d:ɳɳ	137 (11)	34 (9)	150 (19)
	pp:b:m	146 (14)	96 (10)	97 (10)

TABLE 3. Mean closure durations (standard deviations) in ms for geminate and single stops and nasals; the retroflex nasals were all geminate and T_3 's bilabial nasal single. n = at least 10 for each type.

duration for bilabial stops are close to those between stops contrasting for [voice] in English, where [-voice] closures are about 1.5 times the duration of [+voice] ones (Lisker 1957), but are much larger for retroflex stops, where geminate stops have nearly three times the duration of single ones. Balasubramanian's 1980 data suggest an overall ratio of geminate to single consonant durations of about 1.6:1 for Tamil. A burst was detected for all single stops, showing that they were realized neither as fricatives, as is sometimes the case in this language, nor as flaps. Moreover, these data show both consistent duration—note the small standard deviations—and voicing differences. Only the persistence of voicing in single bilabial stop closures was variable; in some tokens voicing died away before the stop was released, while in others it lasted through the entire stop.

All other evidence also points to [length] rather than [voice] as the contrastive feature. First, geminate stops are written in Tamil by doubling the symbol for the single ones. Second, since sonorants differing only in duration contrast in Tamil, pattern congruity favors [length] over [voice] as the contrastive property of stops. Third, a morpheme structure constraint in Tamil requires that stems have either the shape CVVC- or CVCC- (Christdas 1988). Stops as well as sonorants may occur as the geminate postvocalic consonant in CVCC- stems. Fourth, a morphophonemic rule which geminates the final consonant of a stem before certain suffixes applies to stops as well as sonorants.¹⁷

Kingston (1986) also measured F_0 after the stops produced by his three Tamil speakers; mean F_0 contours are presented in Figure 3a–d. For speakers T_2 (Fig. 3b) and T_3 (Fig. 3c–d), the F_0 contours following the two classes of stops are nearly identical, and what differences occur are small both in magnitude and in duration. The small size of these differences supports the claim that F_0 will only differ markedly when vocal fold vibration is contrastive.

However, for speaker T_1 (Fig. 3a), whose stops did not differ otherwise from those of T_2 or T_3 , the F_0 contour following the geminate voiceless stop was slightly but consistently higher than that following the single voiced stop, throughout the following vowel. Though it is doubtful that such persistent, small F_0 differences could come simply from the presence or absence of vocal fold vibration in the preceding consonant, T_1 's data still do not support the claim that F_0 differences are triggered only by contrastive use of the feature [voice].

The very small differences in F_0 observed just after geminate and single stops for two of the Tamil speakers may be automatic byproducts of the presence vs. absence of voicing in the stop closure. The much larger differences found in most languages where the feature [voice] is contrastive might then be an exaggeration of this normally quite small effect. These larger differences are introduced by the phonetic component, rather than phonologically; but, as tone

¹⁷ Lisker & Abramson 1964 demonstrated that initial stops in Tamil are divided by VOT into two categories—prevoiced and voiceless unaspirated—and there is no [length] contrast initially. However, prevoiced stops occur in Tamil only in unassimilated loanwords from Sanskrit (M. Fowler 1954) and are replaced by voiceless unaspirated stops, the only kind that occurs initially in native words, in more colloquial varieties than Lisker & Abramson's speaker produced. These data therefore do not challenge the interpretation of the medial contrast as one of [length] rather than [voice].

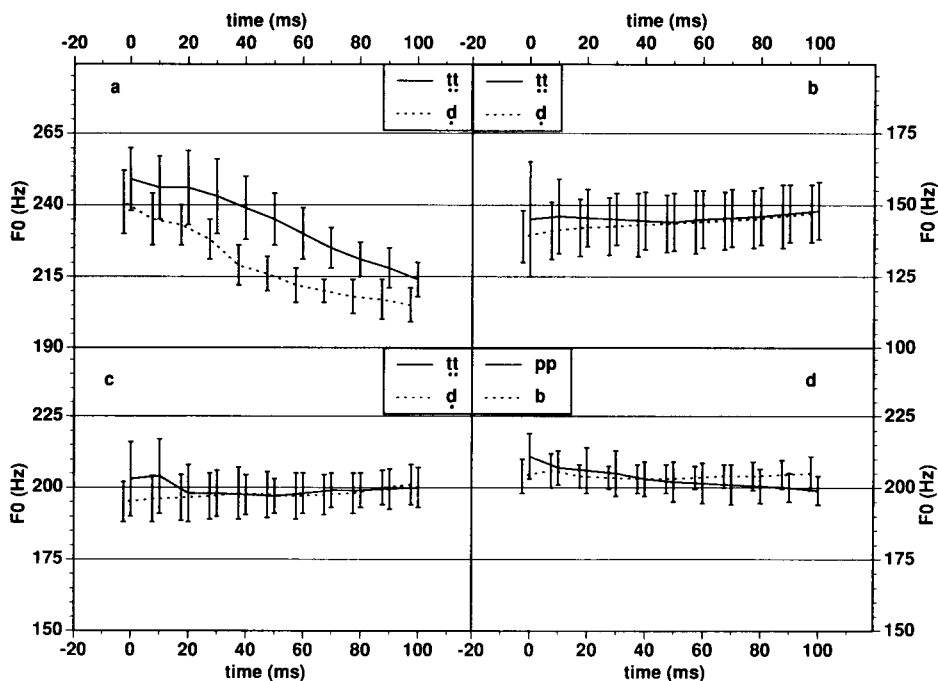


FIGURE 3. F_0 contours (error bars = ± 1 standard deviation) following word-internal, intervocalic [t̪, d̪, pp, b] in nonsense words produced by three adult Tamil speakers (T_1 – T_3). Solid = [t̪t̪, pp], dashed = [d̪d̪, b]. The words were produced in the frame sentence *avan _____ irrenḡu vadavu sonnaan* 'He said _____ again'. The retroflex consonants [t̪, d̪] (a–c) occurred in the context *paa_____al*, while the bilabial consonants [pp, b] (d) occurred in the context *paasi_____a*. n = at least 10 repetitions of each utterance type. F_0 was obtained by measuring period durations. As in the plots in Fig. 1, an F_0 contour was then interpolated at 10 ms intervals before averaging. Subscript dots are used to indicate retroflexion in the figure legends.

splitting and the blocking of tone spreading shows (Hyman & Schuh 1974, Hombert et al. 1979, Kingston & Solnit 1988, 1989; see also Hombert 1977), the phonology can be influenced by them in particular languages.

But what if the F_0 differences after Tamil stops are small because the prosodic context in which they occur is not one which allows such differences to be large? After all, other phonetic differences between contrasting segments may be small or even absent in some prosodic contexts; in particular, supposedly automatic F_0 differences between high and low vowels are small or negligible in German when no (high) pitch accent occurs on the word containing the vowels (Ladd & Silverman 1984). However, F_0 can be noticeably higher next to [–voice] than [+voice] stops even when no high pitch accent occurs on that word (Kingston 1989). Similarly, Silverman 1986 shows that the size of F_0 differences after stops contrasting for [voice] (though not the direction of F_0 change) is insensitive to the larger intonational context in which the stops are

embedded. These facts suggest that [voice]-determined F_0 differences do not depend on the prominence of the word the stops occur in, and therefore that the Tamil stops' prosodic context is not responsible for their negligible effect of F_0 .

3.3. ALTERNATIVE AUTOMATIC MECHANISMS. The foregoing discussion shows that F_0 differences following consonants contrasting for [voice] cannot be predicted from the size of the glottal opening in the consonant or from the occurrence of closure voicing. Rather, they are apparently only reliably predicted from the consonant's phonological specification for [voice]. But what if, whenever a [+voice] or [-voice] consonant is uttered, some other articulation is reliably made that would automatically affect F_0 in the desired ways? Three such mechanisms have been proposed in the literature:

(1) Larynx lowering, which aids the initiation of voicing in initial [+voice] stops by expanding the oral cavity, might also depress F_0 after them (Ewan & Kronen 1974, Hombert et al. 1979, Riordan 1980; see also Ohala 1970, Simada & Hirose 1971, Collier 1974).

(2) Elevated cricothyroid (CT) activity, which aids the extinction of voicing in postvocalic [-voice] stops by increasing horizontal vocal fold tension, might also elevate F_0 in following vowels (Löfqvist et al. 1989; see also Hirose et al. 1972, Hirose et al. 1974, Hirose & Ushijima 1978, Kagaya & Hirose 1975, Collier et al. 1979, Dixit & MacNeilage 1980, Löfqvist et al. 1984).

(3) Elevated transglottal air flow, which is a side effect of an open glottis in [-voice] stops, might also elevate F_0 after them (Kohler 1984, citing data from Barry & Kuenzel 1975, Butcher 1977).

Space limitations prohibit us from discussing these mechanisms in detail, so we will only mention here the strongest arguments against each of them.

Regarding possible effects on F_0 in adjacent vowels of differences in vertical larynx height during stops contrasting for [voice], Kingston 1985 showed that F_0 was at best weakly positively correlated with larynx height in vowels next to [+voice] stops in English and Tigrinya, but uncorrelated with larynx height next to the other kinds of stops in these two languages.

Regarding differences in CT activity, Löfqvist et al.'s 1989 data, as well as those in Löfqvist et al. 1984, show that the amount that F_0 is elevated in the following vowel does not depend on how long the temporal interval is between the end of the preceding vowel when the elevation of CT activity occurs and the beginning of the following vowel. F_0 is elevated as much when the elevated CT activity is further away, as a result of longer intrinsic consonant duration or of a larger number of intervening consonants, as when the CT elevation is closer. Even given a latency of as much as 80 ms between an increase in CT activity and an increase in horizontal fold tension (Collier 1974, Atkinson 1978, Baer 1981; Baer [personal communication, 1989] suggests a much shorter latency, of 20–40 ms), the intervals between when CT becomes active at the end of the preceding vowel and when the following vowel begins appear to be too

long for the contraction of this muscle to be responsible for F_0 elevation in the following vowel in many of the utterances they consider.¹⁸

Regarding differences in transglottal air flow, even if A_g and thus U_g^+ remain relatively large after voicing begins, as observed after the release of German voiceless aspirated stops (Barry & Kuenzel 1975, Butcher 1977), this mechanism implies that the stops' effect on F_0 is a function of glottal aperture: the larger the aperture in the stop, the higher F_0 will be in the following vowel. But we have already seen (in Table 2) that F_0 differences cannot be predicted from the size of the glottal aperture during the preceding stop (see also Hombert & Ladefoged 1976).

Thus, each of these mechanisms is a plausible source of the F_0 differences, but none is free of apparently fatal flaws. It appears unlikely that any aerodynamic mechanism is responsible for the F_0 differences, but we cannot yet rule out other means of changing vocal fold tension than the ones we've discussed. These failures don't rule out some still undiscovered mechanism which is intended to control voicing but which also produces the F_0 differences as an automatic byproduct. However, since the range of possible mechanisms has been reasonably well covered by these three alternatives, the persistence of failure does also invite the alternative view, that the F_0 differences are independently controlled attributes of the [voice] contrast.

The next section lays out what is known about the perceptual interactions among the many acoustic differences between intervocalic stops contrasting for [voice]. We will suggest that lowering F_0 next to a [+voice] stop might act jointly with the stop's other acoustic properties to increase its perceptual distance from a [-voice] stop, and thus may provide a perceptual reason why F_0 should be deliberately lowered in vowels next to a [+voice] stop.

4. LISTENER-ORIENTED PHONETIC KNOWLEDGE. Variation in the realization of the [voice] contrast between initial and intervocalic contexts illustrates a kind of phonetic knowledge that is speaker-oriented in the sense that it concerns aerodynamic and mechanical constraints on speech production. Speakers also use listener-oriented phonetic knowledge when they adapt their articulatory behavior to ensure sufficient auditory distinctiveness of phonological contrasts. Here we explore several types of phonetic redundancy exploited by speakers in signaling phonological contrasts in general and the [voice] contrast in particular (see also Diehl & Kluender 1989a,b, Diehl et al. 1990, Kingston & Diehl 1993, Stevens et al. 1986, Stevens & Keyser 1989).

4.1. THREE TYPES OF PHONETIC REDUNDANCY. Acoustic properties combine in the auditory representation of a distinctive feature value in three distinct ways: (i) each acoustic property may correspond to an independent auditory

¹⁸ Löfqvist et al. (1989) suggest that the latency between CT contraction and F_0 elevation would explain why the preceding vowel's F_0 is not raised as much before a [-voice] stop as a following vowel's F_0 is raised after one. But this would also mean that the increase in horizontal tension brought about by contracting CT would be too late to contribute much to extinguishing voicing at the beginning of the [-voice] stop closure when A_g is still small.

property; (ii) acoustically independent properties may act as subproperties of a single auditory property; and (iii) some subproperties may contribute to more than one auditory property.

4.1.1. MULTIPLE INDEPENDENT CORRELATES. First, there are typically multiple, auditorily independent correlates that serve as distinct bases for a minimal phonological distinction. (These correlates correspond psychologically to what will be referred to below as 'contrastive perceptual properties'.) In the case of the [voice] distinction, one important contrastive property has been identified in Stevens & Blumstein 1981, based on the work of Lisker & Abramson (1964): [+voice] consonants, but not [–voice] consonants, are characterized by the 'presence of low-frequency spectral energy or periodicity over a time interval of 20–30 msec in the vicinity of the acoustic discontinuity that precedes or follows the consonantal constriction interval' (1981:29). We refer to this as the low-frequency property.

In languages such as English, another contrastive property associated with the [voice] distinction (especially initially and pretonically) is the presence or absence of significant aspiration (Repp 1979). Still other contrastive properties are consonant duration and preceding vowel duration: intervocalic [+voice] consonants in most languages have shorter constriction or closure intervals and longer preceding vowels than intervocalic [–voice] consonants. Given this relation between vowel and consonant durations, it is reasonable to try to incorporate both durations into a single measure that may appropriately define an overall durational cue for the [voice] contrast. For example, Kohler (1979) proposed that the distinction between intervocalic fortis and lenis consonants—his terms for what we call [–voice] and [+voice]—is well specified by the duration ratio, vowel/(vowel + consonant). In a related proposal for the English intervocalic [voice] distinction, Port & Dalby (1982) suggested that the consonant/vowel duration ratio is perceptually the most relevant cue (cf. Massaro & Cohen 1983). We will refer to this contrastive perceptual property as the C/V duration ratio.

The presence or absence of the low-frequency property, the presence or absence of aspiration, and the C/V duration ratio each correspond to an auditorily independent perceptual variable, and collectively they provide redundant specification of the [voice] contrast. Although not every such variable is exploited in every language or in every utterance position, typically more than one contrastive perceptual property will signal that a consonant is [+voice] or [–voice]. Thus, in English initial consonants, the [voice] contrast is signaled both by the presence or absence of the low-frequency property and by the presence or absence of aspiration, whereas in intervocalic consonants the contrast is signaled by the presence or absence of the low-frequency property and by the C/V duration ratio.

4.1.2. MULTIPLE SUBPROPERTIES. Second, each of the auditorily independent properties of a phonological distinction may typically be analyzed into multiple subproperties that are mutually enhancing in the sense that they all contribute to the same contrastive perceptual property. Consider, for example, the low-

frequency property of [+voice] consonants. As Stevens & Blumstein (1981) pointed out, this property can be analyzed into at least three phonetically separate subproperties—voicing during the consonant constriction interval, a low F_1 near the constriction interval, and a low F_0 in the same region. All three of these subproperties are typical correlates of [+voice] consonants, and each has been found to cue or enhance the perception of the [+voice] category (Haggard et al. 1970, Fujimura 1971, Lisker 1975, 1986, Summerfield & Haggard 1977). The perceptual evidence to date is consistent with Stevens & Blumstein's claim that the three subproperties integrate into a single contrastive perceptual property.

A similar kind of analysis may also be applied to the C/V duration ratio associated with the [voice] contrast. It is obvious that the difference between [+voice] and [–voice] consonants in C/V duration ratio may be enhanced by varying either the consonant duration or the vowel duration, or both. Just as voicing during closure, a low F_1 , and a low F_0 are mutually enhancing in that all contribute to the low-frequency property, a short consonant and a long preceding vowel are mutually enhancing in that they both contribute to the relatively small C/V duration ratio characteristic of [+voice] consonants. Relative to either durational cue in isolation, a ratio of the two durations permits a considerably wider range of variation and hence greater potential distinctiveness.

For a claim that contrastive perceptual properties consist of multiple subproperties that are mutually enhancing, it is important to show that the enhancement effect arises from the general auditory character of the subproperties rather than from their experienced covariation in speech. For example, the cue value of voicing and a low F_1 might result simply from the fact that they are both correlates of [+voice] consonants, otherwise lacking the kind of auditory commonality that gives rise to an integrated contrastive perceptual property.

One way to evaluate the strong claim of auditory enhancement is to examine how two properties such as voicing and a low F_1 interact perceptually when the stimuli that contain them are not perceived phonetically. Kingston & Diehl (1993) had listeners classify single-formant analogues of /aba/–/apa/ stimuli in which presence or absence of laryngeal pulsing during the medial gap was varied orthogonally with the F_1 offset value at the edge of the gap. Although the stimuli were not speech-like (and therefore the effects of linguistic experience would presumably be irrelevant), voicing and F_1 offset exhibited the predicted enhancement relation. In particular, the most reliable classification was obtained for the stimulus pair in which voicing co-occurred with a low F_1 offset and absence of voicing co-occurred with a high F_1 offset. When these relations were reversed, success in classifying the stimuli was significantly reduced. These results suggest that voicing and a low F_1 do indeed contribute to a single contrastive perceptual property, i.e. the low-frequency property.

4.1.3. MULTIPLE USES OF SUBPROPERTIES. Third, the subproperties that contribute to one contrastive perceptual property of a phonological distinction also

tend to contribute to other contrastive perceptual properties. That is, they have more than one perceptual role and these roles may be auditorily independent.

Lisker 1957 showed that variation in closure duration is sufficient to signal the distinction between intervocalic [+voice] and [–voice] consonants. Later, Lisker (1986 [1978]) found that the presence of voicing during the closure yields an increase in [+voice] identification responses. Parker et al. (1986) conducted a modified version of the latter study by Lisker. Two stimulus series ranging perceptually from /aba/ to /apa/ were created by varying the closure duration of the consonant. The two series differed only with respect to the presence or absence of laryngeal pulsing during the closure. As expected, variation in closure duration proved to be sufficient to cue the /b/–/p/ distinction, and the presence of voicing during closure shifted the /b/–/p/ identification boundary toward larger values of closure duration (that is, there were more [+voice] responses).

The boundary shift caused by the presence of pulsing during closure is, of course, predicted by the fact that it contributes to the low-frequency property which is characteristic of [+voice] consonants. However, Parker et al. hypothesized that the presence of pulsing has another effect as well—namely, it reduces the perceived closure duration, making the stimulus appear even more strongly [+voice].

To test this hypothesis, Parker et al. also prepared three sets of nonspeech analogues of the /aba/–/apa/ stimuli, consisting of square-wave segments separated by a medial gap of varying duration. The gap contained either silence or else the same segment of pulsing used in the corresponding speech stimuli. The three nonspeech stimulus sets differed only in the F_0 trajectory in the vicinity of the medial gap: F_0 remained constant, rose before and fell after the gap, or fell before and rose after the gap. After training with feedback on the endpoints for gap duration, subjects judged each item in the series on the basis of similarity to the short-gap or long-gap training stimuli.

Analogous to the results for the /aba/–/apa/ stimuli, the presence of pulsing during the medial gap produced a significant category boundary shift, but only for the stimulus set in which F_0 fell before and rose after the gap. This shift was in the same direction as that for the /aba/–/apa/ stimuli, but only about one third the magnitude. Thus, at least in the one stimulus condition, the presence of pulsing made the gap between the square-wave segments appear smaller in duration. This is consistent with the hypothesis of Parker et al. that one effect of voicing is to enhance the closure-duration cue for [+voice] stops.

What accounts for the differences across the various speech and nonspeech conditions in the size of the boundary shift induced by the presence of pulsing? The relatively large boundary shift in the /aba/–/apa/ condition can be explained on the assumption that voicing serves at least two independent perceptual roles in specifying the [+voice] category. First, it is a major contributor to the low-frequency property, a main contrastive perceptual property of [+voice] consonants. Second, it enhances the closure-duration cue (or C/V duration ratio cue) by making brief closures seem even shorter. However, in the square-wave

conditions, pulsing during the intervocalic gap serves, at most, only the latter of these two roles. This is because the nonspeech training categories were defined on the basis of gap duration alone and were uncorrelated with the presence or absence of pulsing. Thus, for the square-wave categories, the low-frequency property per se has no distinctive role, and the effect of pulsing is limited to altering the apparent duration of the intervocalic gap.

What remains to be accounted for is why, among the three square-wave conditions, pulsing had a significant effect only in the fall-rise condition. A tentative explanation is that a falling pattern before the gap and a rising pattern after the gap make the pulsing more continuous spectrally with the flanking sounds, and that this continuity is necessary for pulsing to be integrated with the rest of the signal so as to influence the perceived duration of the gap. Perceptual integration of temporally distinct signals has been shown to be enhanced by spectral continuity (e.g. Bregman & Dannenbring 1973).

In view of the pattern of nonspeech findings, it is noteworthy that a falling-rising spectral pattern is characteristic of vowels flanking [+voice] stops in natural speech. As described earlier, both F_1 and F_0 have relatively low values in the vicinity of [+voice] consonants, thus contributing to the integrated percept we have labeled the low-frequency property. However, the square-wave results suggest that these parameters may also influence intervocalic [voice] judgments in a quite independent way, namely by creating a higher degree of spectral continuity between closure pulsing and the flanking vowels, thus contributing to a perceived shortening of the closure.

To evaluate this last claim, Kingston et al. (1990) replicated the nonspeech experiment of Parker et al. 1986, using single formant stimuli in which F_1 and F_0 could be independently manipulated. All possible patterns of steady-state, falling-rising, and rising-falling F_1 and F_0 trajectories were combined with gaps of varying duration, with and without pulsing. For all three F_1 contours, the presence of pulsing increased the percentage of short-gap responses, but the effect was largest when F_1 fell before and rose after the medial gap. The effect of pulsing on perceived gap duration was independent of the F_0 contour, however. Thus, the perceived shortening of a closure interval caused by the presence of voicing is enhanced by spectral continuity between the voicing segment and the flanking vowels, and, moreover, this spectral continuity effect rests on the F_1 contour rather than on the F_0 contour. We conclude that a low F_1 not only contributes to the low-frequency property of [+voice] consonants, but also enhances the C/V duration ratio by serving as a prerequisite for voicing to reduce significantly the apparent duration of the closure. A low F_0 , by contrast, contributes only to the low-frequency property of [+voice] consonants.

The interlocking network of mutual enhancement relations apparently does not end there. Javkin 1976 reported that, when listeners were asked to vary the duration of a tone to match that of the vowel, the presence of voicing in the following consonant yielded tone durational settings that were reliably longer. This means that voicing contributes to a more distinctive C/V duration ratio in two ways: by shortening the apparent duration of the consonant and by lengthening the apparent duration of the vowel. These effects, together with

the contribution of voicing to the low-frequency property, help to explain the perceptual robustness of the [voice] contrast.

Our general claim is that many regularities of phonetic covariation reflect a strategy used by speakers to enhance the auditory distinctiveness of phonological contrasts. An alternative view is that most correlates of, for example, the [voice] contrast are simply physical or physiological byproducts of controlling the voicing variable. In §3, we argued that this interpretation of the F_0 correlate of the [voice] contrast is incorrect. Analogous arguments with respect to the closure duration and vowel duration correlates of the [voice] contrast are discussed by Lisker (1974), Javkin (1976), and Kluender et al. (1988).

The next section elaborates upon the role of the listener-oriented phonetic knowledge just described in mediating the phonetic realization of phonological contrasts.

4.2. AUDITORY CONSTRAINTS AND LISTENER-ORIENTED PHONETIC KNOWLEDGE. In §§2–3 the argument for phonetic knowledge relied heavily on the proposition that regularities of phonetic implementation are motivated—but not uniquely determined—by physical constraints on production. The articulatory covariation used to signal intervocalic [voice] contrasts is neither automatic nor fully explained by auditory constraints. Rather, this covariation is favored crosslinguistically because it contributes to a robust perceptual contrast. The preceding discussion described three aspects of this perceptual robustness in intervocalic stops. The covariation in initial stops was explained by appealing to speaker-oriented phonetic knowledge, specifically the relatively great articulatory cost of initiating voicing, while the regularities in intervocalic stops can be explained by appealing to listener-oriented phonetic knowledge, specifically the perceptual robustness of acoustic effects that integrate into higher-level perceptual properties.

Like the laws of physics, certain aspects of auditory perception are automatic in the sense that they cannot be directly altered or controlled by human intentions. For example, we assume that the perceived shortening of a closure interval induced by the presence of voicing occurs automatically as a result of normal auditory processes. How, then, is phonetic knowledge involved? For the speaker, a listener-oriented strategy requires knowing, among other things, that voicing has this perceptual shortening effect and can therefore be exploited to enhance the closure-duration cue for [voice] contrasts. For the listener, phonetic knowledge embraces such facts as that English intervocalic [voice] contrasts are specified by the presence or absence of the low-frequency property and by a ratio of the perceived durations of the consonant and preceding vowel. What properties and dimensions count as perceptually relevant to a phonological contrast is therefore by no means automatic, even if some of the auditory processes that affect the perception of these dimensions are.

The preceding discussion suggests that certain patterns of covariation may be far more easily explained with respect to these general auditory effects than with respect to articulation; that is, one acoustic dimension covaries with another because of their perceptual interaction rather than because of an articula-

tory dependency. This perspective implies that speakers have knowledge of the mechanisms that listeners apply to the task of recognizing speech sounds, and that this knowledge prescribes reorganizations of articulatory behaviors to take advantage of these mechanisms. It should be clear that such reorganizations require the phonetic component to be controlled.

5. CONCLUSION. We have attempted to show that the ways in which speech sounds are realized are not determined entirely by their phonological specification for distinctive features or by constraints imposed by the speech production or perception apparatus. Instead, we have argued, in implementing phonological contrasts a complex intermediate device is needed to control sets of articulations. This control limits energy expenditures or produces arrays of mutually enhancing acoustic effects. We have also shown that the typical output of this device—at least for [voice] contrasts—is distinct phonetic categories, rather than continuous variation along phonetic dimensions.

REFERENCES

- ABBS, JAMES H.; VINCENT L. GRACCO; and KELLY J. COLE. 1984. Control of multi-movement coordination: Sensory motor mechanisms in speech motor programming. *Journal of Motor Behavior* 16.195–231.
- ATKINSON, JAMES E. 1978. Correlation analysis of the physiological features controlling voice fundamental frequency. *Journal of the Acoustical Society of America* 63.211–22.
- BAER, THOMAS. 1981. Investigation of the phonatory mechanism. Proceedings of the conference on the assessment of vocal pathology, ASHA Reports 11.38–46.
- BALASUBRAMANIAN, T. 1980. Timing in Tamil. *Journal of Phonetics* 8.449–67.
- BARRY, WILLIAM J., and HERMANN J. KUENZEL. 1975. Co-articulatory airflow characteristics of intervocalic voiceless plosives. *Journal of Phonetics* 3.263–82.
- BENQUEREL, ANDRÉ-PIERRE, and TEJ BHATIA. 1980. Hindi stop consonants: An acoustic and fiberoptic study. *Phonetica* 37.134–48.
- BREGMAN, ALBERT S., and GARY L. DANNENBRING. 1973. The effect of continuity on auditory stream segregation. *Perception & Psychophysics* 13.308–12.
- BROWMAN, CATHERINE P., and LOUIS M. GOLDSTEIN. 1990. Tiers in articulatory phonology, with some implications for casual speech. Papers in laboratory phonology I: Between the grammar and physics of speech, ed. by John Kingston and Mary E. Beckman, 341–76. Cambridge: Cambridge University Press.
- . 1992. Articulatory phonology: An overview. *Phonetica* 49.155–80.
- BUTCHER, ANTHONY. 1977. Coarticulation in intervocalic voiceless plosives and fricatives in connected speech. Kiel: Universität Kiel, Institut für Phonetik, Arbeitsberichte 8.154–212.
- CAISSE, MICHELLE. 1982. Cross-linguistic differences in fundamental frequency perturbation induced by voiceless unaspirated stops. Berkeley, CA: University of California M.A. thesis.
- CHEN, MATTHEW. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22.129–59.
- CHOMSKY, NOAM, and MORRIS HALLE. 1968. The sound pattern of English. New York: Harper and Row.
- CHRISTIDAS, PRATHIMA. 1988. Tamil morphology and phonology. Ithaca, NY: Cornell University dissertation.
- COLEMAN, JOHN S. 1992. The phonetic interpretation of headed phonological structures containing overlapping constituents. *Phonology* 9.1–44.

- COLLIER, RENÉ. 1974. Laryngeal muscle activity, subglottal air pressure, and the control of pitch in speech. New Haven, CT: Haskins Laboratories Status Report on Speech Research SR-39/40.137-70.
- ; LEIGH LISKER; HAJIME HIROSE; and TATSUJIRO USHIJIMA. 1979. Voicing in intervocalic stops and fricatives in Dutch. *Journal of Phonetics* 7.357-73.
- COOPER, ANDRÉ M. 1991a. Glottal gestures and aspiration in English. New Haven, CT: Yale University dissertation.
- . 1991b. Laryngeal and oral gestures in English /p,t,k/. Aix-en-Provence: Proceedings of the XIIth International Congress of Phonetic Sciences, vol. 2.50-53.
- DAVIS, STUART, and W. VAN SUMMERS. 1989. Vowel length and closure duration in word-medial VC sequences. *Journal of Phonetics* 17.339-53.
- DIEHL, RANDY L., and KEITH R. KLUENDER. 1989a. On the objects of speech perception. *Ecological Psychology* 1.121-44.
- , —. 1989b. Reply to commentators. *Ecological Psychology* 1.195-225.
- ; —; and MARGARET A. WALSH. 1990. Some auditory bases of speech perception and production. *Advances in speech, hearing and language processing*, vol. 1, ed. by William A. Ainsworth, 243-67. London: JAI Press.
- DIXIT, R. PRAKASH. 1989. Glottal gestures in Hindi plosives. *Journal of Phonetics* 17.213-37.
- , and PETER F. MACNEILAGE. 1980. Cricothyroid activity and the control of voicing in Hindi stops and affricates. *Phonetica* 37.397-406.
- DOCHERTY, GERARD J. 1989. An experimental phonetic study of the timing of voicing in English obstruents. Edinburgh: University of Edinburgh dissertation.
- ERICKSON, DONNA. 1975. Phonetical implications for a historical account of tonogenesis in Thai. *Studies in Thai linguistics in honor of William J. Gedney*, ed. by Jimmy Harris and James R. Chamberlin, 100-111. Bangkok: Central Institute of the English Language.
- EWAN, WILLIAM G. 1976. Laryngeal behavior in speech. Berkeley, CA: University of California dissertation. [Reprinted 1979, University of California, Berkeley, Report of the Phonology Laboratory 3.]
- . 1979. Can vowel intrinsic F_0 be explained by source/tract coupling? *Journal of the Acoustical Society of America* 66.358-62.
- , and ROBERT KRONES. 1974. Measuring larynx movement using the thyroumbrometer. *Journal of Phonetics* 2.327-35.
- FISCHER-JØRGENSEN, ELI. 1968. Voicing, tenseness and aspiration in stop consonants, with special reference to French and Danish. Copenhagen: Annual Report of the Institute of Phonetics, University of Copenhagen 3.63-114.
- . 1985. Some basic vowel features, their articulatory correlates, and their explanatory power in phonology. *Phonetic linguistics: Essays in honor of Peter Ladefoged*, ed. by Victoria A. Fromkin, 79-99. New York: Academic Press.
- . 1990. Intrinsic F_0 in tense and lax vowels with special reference to German. *Phonetica* 47.99-140.
- FLEGE, JAMES E. 1982. Laryngeal timing and phonation onset in utterance-initial stops. *Journal of Phonetics* 10.177-92.
- FOWLER, CAROL A. 1990. Some regularities in speech are not consequences of formal rules: Comments on Keating's paper. *Papers in laboratory phonology I: Between the grammar and physics of speech*, ed. by John Kingston and Mary E. Beckman, 476-89. Cambridge: Cambridge University Press.
- , and MICHAEL T. TURVEY. 1980. Immediate compensation in bite-block speech. *Phonetica* 37.306-26.
- ; PHILIP RUBIN; ROBERT E. REMEZ; and MICHAEL T. TURVEY. 1980. Implications for speech production of a general theory of action. *Language production I: Speech and talk*, ed. by Brian Butterworth, 372-420. London: Academic Press.
- FOWLER, MURRAY. 1954. The segmental phonemes of Sanskritized Tamil. *Lg.* 30.360-67.

- FOX, ROBERT ALLEN, and DALE TERBEEK. 1977. Dental flaps, vowel duration and rule ordering. *Journal of Phonetics* 5.27–34.
- FRØKJAER-JENSEN, BØRGE; CARL LUDVIGSEN; and JØRGEN RISCHÉL. 1971. A glotto-graphic study of some Danish consonants. Form and substance: Phonetic and linguistic papers presented to Eli Fischer-Jørgensen, ed. by L. L. Hammerich, Roman Jakobson, and Eberhard Zwirner, 123–40. Copenhagen: Akademisk Forlag.
- FUJIMURA, OSAMU. 1971. Remarks on stop consonants: Synthesis experiments and acoustic cues. Form and substance: Phonetic and linguistic papers presented to Eli Fischer-Jørgensen, ed. by L. L. Hammerich, Roman Jakobson and Eberhard Zwirner, 221–32. Copenhagen: Akademisk Forlag.
- FUKUI, N., and HAJIME HIROSE. 1983. Laryngeal adjustments in Danish voiceless obstruent production. Tokyo: University of Tokyo Research Institute in Logopedics and Phoniatrics, Annual Bulletin 17.61–71.
- GANDOUR, JACKSON. 1974. Consonant types and tone in Siamese. *Journal of Phonetics* 2.337–50.
- HAGGARD, MARK P.; STEPHEN AMBLER; and MO CALLOW. 1970. Pitch as a voicing cue. *Journal of the Acoustical Society of America* 47.613–17.
- HALLE, MORRIS. 1983. On distinctive features and their articulatory implementaton. *Natural Language and Linguistic Theory* 1.91–105.
- HAN, MEIKO S., and RAYMOND S. WEITZMAN. 1970. Acoustic features of Korean /P,T,K/, /p,t,k/, and /p^h,t^h,k^h/. *Phonetica* 22.112–28.
- HARDCASTLE, WILLIAM J. 1973. Some observations on the tense-lax distinction in initial stops in Korean. *Journal of Phonetics* 1.263–72.
- HIROSE, HAJIME, and TATSUJIRO USHIJIMA. 1978. Laryngeal control for the voicing distinction in Japanese. *Phonetica* 35.1–10.
- HIROSE, HAJIME; C. Y. LEE; and TATSUJIRO USHIJIMA. 1974. Laryngeal control in Korean stop production. *Journal of Phonetics* 2.145–52.
- HIROSE, HAJIME; LEIGH LISKER; and ARTHUR S. ABRAMSON. 1972. Physiological aspects of certain laryngeal features in stop production. New Haven, CT: Haskins Laboratories Status Report on Speech Research SR-31/32.183–91.
- HIROSE, HAJIME; SEIJI NIIMI; KIYOSHI HONDA; and MASAYUKI SAWASHIMA. 1985. The relationship between glottal opening and the transglottal pressure difference during consonant production. Tokyo: University of Tokyo Research Institute in Logopedics and Phoniatrics, Annual Bulletin 19.55–64.
- HIROSE, HAJIME; H. S. PARK; HIROHIDE YOSHIOKA; MASAYUKI SAWASHIMA; and H. UMEMA. 1981. An electromyographic study of laryngeal adjustments for the Korean stops. Tokyo: University of Tokyo Research Institute in Logopedics and Phoniatrics, Annual Bulletin 15.31–43.
- HOMBERT, JEAN-MARIE. 1977. Consonant types, vowel height and tone in Yoruba. *Studies in African Linguistics* 8.173–90.
- . 1978. Consonant types, vowel quality, and tone. *Tone: A linguistic survey*, ed. by Victoria A. Fromkin, 77–111. New York, NY: Academic Press.
- , and PETER LADEFOGED. 1976. The effect of aspiration on the fundamental frequency of the following vowel. *Journal of the Acoustical Society of America* 59.S72 (abstract).
- ; JOHN J. OHALA; and WILLIAM G. EWAN. 1979. Phonetic explanations for the development of tones. *Lg.* 55.37–58.
- HONDA, KIYOSHI. 1983. Relationship between pitch control and vowel articulation. *Vocal fold physiology: Contemporary research and clinical issues*, ed. by Diane M. Bless and James H. Abbs, 286–97. San Diego, CA: College-Hill Press.
- , and OSAMU FUJIMURA. 1991. Intrinsic vowel F₀ and phrase-final lowering: Phonological vs. biological explanations. *Vocal fold physiology: Acoustic, perceptual, and physiological aspects of voice mechanisms*, ed. by Jan Gauffin and Britta Hammerberg, 149–57. San Diego, CA: Singular Publishing Group.

- HUTTERS, BIRGIT ELISABETH. 1985. Vocal fold adjustments in aspirated and unaspirated stops in Danish. *Phonetica* 42.1–24.
- HYMAN, LARRY M., and RUSSELL G. SCHUH. 1974. Universals of tone rules: Evidence from West Africa. *Linguistic Inquiry* 5.81–115.
- IWATA, RAY, and HAJIME HIROSE. 1976. Fiberoptic acoustic studies of Mandarin stops and affricates. Tokyo: University of Tokyo Research Institute in Logopedics and Phoniatics, *Annual Bulletin* 10.47–60.
- IWATA, RAY; MASAYUKI SAWASHIMA; and HAJIME HIROSE. 1981. Laryngeal adjustments for syllable-final stops in Cantonese. Tokyo: University of Tokyo Research Institute in Logopedics and Phoniatics, *Annual Bulletin* 15.45–54.
- ; —; —; and SEJI NIIMI. 1979. Laryngeal adjustments of Fukienese stops—Initial plosives and final aplosives. Tokyo: University of Tokyo Research Institute in Logopedics and Phoniatics, *Annual Bulletin* 13.61–81.
- JAEGER, JERI J. 1978. Speech aerodynamics and phonological universals. *Berkeley Linguistics Society* 4.311–29.
- JAVKIN, HECTOR R. 1976. The perceptual basis of vowel duration differences associated with the voiced/voiceless distinction. Berkeley: University of California, Report of the Phonology Laboratory 1.78–92.
- JEEL, V. 1975. An investigation of the fundamental frequency of vowels after various Danish consonants, in particular stop consonants. Copenhagen: Annual Report of the Institute of Phonetics, University of Copenhagen 9.191–211.
- JUN, SUN-AH. 1992. The status of lenis stop voicing rule in Korean. Paper presented at the 8th International Conference on Korean Linguistics. Proceedings of the student conference in linguistics, Stanford, CA: Stanford University Press, to appear.
- KAGAYA, RYOHEI. 1971. Laryngeal gesture in Korean stop consonants. Tokyo: University of Tokyo Institute of Logopedics and Phoniatics, *Annual Bulletin* 5.15–23.
- . 1974. A fiberoptic and acoustic study of Korean stops, affricates, and fricatives. *Journal of Phonetics* 2.161–80.
- , and HAJIME HIROSE. 1975. Fiberoptic, electromyographic, and acoustic analyses of Hindi stop consonants. Tokyo: University of Tokyo Institute of Logopedics and Phoniatics, *Annual Bulletin* 9.27–46.
- KAWASAKI, HARUKO. 1983. Fundamental frequency perturbation caused by voiced and voiceless stops in Japanese. Cambridge, MA: MIT Research Laboratory of Electronics, Speech Communication Group, Working Papers 3.55–67.
- KEATING, PATRICIA A. 1984. Phonetic and phonological representation of stop consonant voicing. *Lg.* 60.286–319.
- . 1985. Universal phonetics and the organization of grammars. *Phonetic linguistics: Essays in honor of Peter Ladefoged*, ed. by Victoria A. Fromkin, 115–32. Orlando, FL: Academic Press.
- . 1988. The phonology-phonetics interface. *Linguistics: The Cambridge survey*, vol. 1, ed. by Frederick J. Newmeyer, 281–302. Cambridge: Cambridge University Press.
- . 1990. Phonetic representations in a generative grammar. *Journal of Phonetics* 18.321–34.
- ; WENDY LINKER; and MARIE K. HUFFMAN. 1983. Patterns in allophonic distribution for voiced and voiceless stops. *Journal of Phonetics* 11.277–90.
- KELSO, J. A. SCOTT, and BETTY TULLER. 1983. "Compensatory articulation" under conditions of reduced afferent information: A dynamic formulation. *Journal of Speech and Hearing Research* 26.217–24.
- KIM, CHIN-WU. 1965. On the autonomy of the tensivity feature in stop classification (with special reference to Korean stops). *Word* 21.339–59.
- . 1970. A theory of aspiration. *Phonetica* 21.107–16.
- KINGSTON, JOHN. 1985. The linguistic use of vertical larynx movement. Austin: University of Texas, MS. [Expansion of 'The ineffectiveness of vertical larynx movement', *Journal of the Acoustical Society of America* 77.S86–87, 1985 (abstract).]

- . 1986. Are F_0 differences after stops deliberate or accidental? *Journal of the Acoustical Society of America* 79.S27 (abstract).
- . 1989. The effect of macroprosodic context on consonantal perturbations of fundamental frequency. *Journal of the Acoustical Society of America* 85.S149 (abstract).
- . 1991. Integration of articulations in the perception of vowel height. *Phonetica* 47.149–79.
- , and RANDY L. DIEHL. 1993. Intermediate properties in the perception of distinctive feature values. *Papers in laboratory phonology IV*, ed. by Bruce Connell and Amalia Arvaniti. Cambridge: Cambridge University Press, to appear.
- ; —; KEITH R. KLUENDER; and ELLEN M. PARKER. 1990. Resonance versus source characteristics in perceiving spectral continuity between vowels and consonants. *Journal of the Acoustical Society of America* 88.S54–S55 (abstract).
- KINGSTON, JOHN, and DAVID SOLNIT. 1988. *The tones of consonants*. Amherst & Ann Arbor: University of Massachusetts & University of Michigan, ms.
- , —. 1989. The inadequacy of underspecification. *North Eastern Linguistic Society* 19.264–278.
- KLUENDER, KEITH R.; RANDY L. DIEHL; and BEVERLY A. WRIGHT. 1988. Vowel-length differences after voiced and voiceless consonants: An auditory explanation. *Journal of Phonetics* 16.153–69.
- KOHLER, KLAUS J. 1979. Dimensions in the perception of fortis and lenis plosives. *Phonetica* 36.332–43.
- . 1982. F_0 in the production of lenis and fortis plosives. *Phonetica* 39.199–218.
- . 1984. Phonetic explanation in phonology. The feature fortis/lenis. *Phonetica* 41.150–74.
- . 1985. F_0 in the perception of lenis and fortis plosives. *Journal of the Acoustical Society of America* 78.21–32.
- LADD, D. ROBERT, and KIM E. A. SILVERMAN. 1984. Intrinsic pitch of vowels in connected speech. *Phonetica* 41.31–40.
- LADEFOGED, PETER. 1980. What are linguistic sounds made of? *Lg.* 56.485–502.
- ; JOSEPH L. DECLERK; MONA LINDAU; and GEORGE PAPÇUN. 1972. *An auditory-motor theory of speech production*. Los Angeles: UCLA Working Papers in Phonetics 22.48–75.
- LAEUFER, CHRISTIANE. 1992. Patterns of voicing-conditioned vowel duration in French and English. *Journal of Phonetics* 20.411–40.
- LINDBLÖM, BJÖRN. 1983. Economy of speech gestures. *The production of speech*, ed. by Peter MacNeilage, 217–46. Berlin: Springer.
- . 1990. Explaining phonetic variation: A sketch of the H & H theory. *Speech production and speech modeling*, ed. by William J. Hardcastle and Alain Marchal, 403–39. Dordrecht: Kluwer.
- ; JAMES F. LUBKER; and THOMAS GAY. 1979. Formant frequencies of some fixed mandible vowels and a model of predictive simulation. *Journal of Phonetics* 7.147–62.
- LINDQVIST, J. 1972. *Laryngeal articulation studies in Swedish subjects*. Stockholm: Royal Institute of Technology Speech Transmission Laboratory, Quarterly Progress and Status Report 1972/2–3.10–27.
- LISKER, LEIGH. 1957. Closure duration and the intervocalic voiced-voiceless distinctions in English. *Lg.* 33.42–49.
- . 1958. The Tamil occlusives: Short vs. long or voiced vs. voiceless? *Indian Linguistics* (Turner Jubilee Volume I) 19.294–301.
- . 1974. On 'explaining' vowel duration variation. *Glossa* 8.223–46.
- . 1975. Is it VOT or a first formant transition detector? *Journal of the Acoustical Society of America* 57.1547–51.
- . 1986. "Voicing" in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech* 29.3–11. [Also appeared in 1978 as Rapid vs.

- rabid: A catalogue of cues, New Haven, CT: Haskins Laboratories Status Report on Speech Research SR-54.127-132, and *Journal of the Acoustical Society of America* 62.S77 (abstract).]
- , and ARTHUR S. ABRAMSON. 1964. A cross-linguistic study of voicing in initial stops: Acoustical measurements. *Word* 20.384-422.
- , —. 1971. Distinctive features and laryngeal control. *Lg.* 47.767-85.
- LISKER, LEIGH, and THOMAS BAER. 1984. Laryngeal management at utterance-internal word boundary in American English. *Language and Speech* 27.163-71.
- LOFQVIST, ANDERS. 1975a. Intrinsic and extrinsic F_0 variations in Swedish tonal accents. *Phonetica* 31.228-47.
- , —. 1975b. A study of subglottal pressure during the production of Swedish stops. *Journal of Phonetics* 3.175-89.
- , —. 1976. Closure duration and aspiration for Swedish stops. Lund: Lund University Department of General Linguistics, Phonetics Laboratory, Working Papers 13.1-39.
- , —. 1980. Interarticulator programming in stop production. *Journal of Phonetics* 8.475-90.
- , —. 1986. Stability and change. *Journal of Phonetics* 14.139-44.
- , —. 1990. Speech as audible gestures. *Speech production and speech modeling*, ed. by William J. Hardcastle and Alain Marchal, 289-322. Dordrecht: Kluwer.
- , and HIROHIDE YOSHIOKA. 1980. Laryngeal activity in Swedish obstruent clusters. *Journal of the Acoustical Society of America* 68.792-801.
- , —. 1981a. Interarticulator programming in obstruent production. *Phonetica* 38.21-34.
- , —. 1981b. Laryngeal activity in Icelandic obstruent production. *Nordic Journal of Linguistics* 4.1-18.
- , —. 1984. Intra-segmental timing: Laryngeal-oral coordination in voiceless consonant production. *Speech Communication* 3.279-89.
- LOFQVIST, ANDERS; NANCY S. MCGARR; and KIYOSHI HONDA. 1984. Laryngeal muscles and articulatory control. *Journal of the Acoustical Society of America* 76.951-54.
- LOFQVIST, ANDERS; THOMAS BAER; NANCY S. MCGARR; and ROBIN SEIDER STORY. 1989. The cricothyroid muscle in voicing control. *Journal of the Acoustical Society of America* 85.1314-21.
- MACK, MOLLY. 1982. Voicing-dependent vowel duration in English and French: Monolingual vs. bilingual production. *Journal of the Acoustical Society of America* 71.173-78.
- MACNEILAGE, PETER F. 1970. Motor control of serial ordering of speech. *Psychological Review* 77.182-96.
- MASSARO, DOMINIC W., and MICHAEL M. COHEN. 1983. Consonant/vowel ratios: An improbable cue in speech. *Perception & Psychophysics* 33.501-505.
- MENN, LISE. 1983. Development of articulatory, phonetic, and phonological capabilities. *Language production*, vol. 2, ed. by Brian Butterworth, 3-50. New York, NY: Academic Press.
- MONTAIGNE, MICHEL DE. 1587-1588 [1957]. *Of experience*. Complete works: Essays, travel journal, letters. 1533-1592, trans. by Donald M. Frame, 815-857. Stanford, CA: Stanford University Press.
- MULLER, ERIC M., and W. S. BROWN, JR. 1980. Variations in the supraglottal waveform and their articulatory interpretation. *Speech and language: Advances in basic research and practice* 4, ed. by Norman J. Lass, 317-89. New York: Academic Press.
- MUNHALL, KEVIN G., and ANDERS LOFQVIST. 1992. Gestural aggregation in speech. *Journal of Phonetics* 20.111-26.
- NEAREY, TERRANCE M. 1978. *Features for vowels*. Storrs, CT: University of Connecticut dissertation.
- NIHILANI, PAROO. 1974. An aerodynamic study of stops in Sindhi. *Phonetica* 29.193-224.

- NITTROUER, SUSAN. 1992. Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. *Journal of Phonetics* 20.351–82.
- OHALA, JOHN J. 1970. Aspects of the control and production of speech. Los Angeles: UCLA Working Papers in Phonetics 15.
- . 1981a. The listener as a source of sound change. Papers from the parasession on language and behavior, ed. by Carrie S. Masek, Roberta A. Hendrick, and Mary Frances Miller, 178–203. Chicago, IL: Chicago Linguistic Society.
- . 1981b. Articulatory constraints on the cognitive representation of speech. The cognitive representation of speech, ed. by Thomas Myers, John Laver, and John Anderson, 111–22. Amsterdam: North-Holland.
- , and BRIAN M. EUKEL. 1987. Explaining the intrinsic pitch of vowels. In honour of Ilse Lehiste, ed. by Robert Channon and Linda Shockey, 207–15. Dordrecht: Foris.
- , and CAROL J. RIORDAN. 1979. Passive vocal tract enlargement during voiced stops. Speech communication papers from the 97th meeting of the Acoustical Society of America, ed. by Jared Wolf and Dennis H. Klatt, 89–92.
- OHDE, RALPH N. 1984. Fundamental frequency as an acoustic correlate of stop consonant voicing. *Journal of the Acoustical Society of America* 75.224–30.
- PARKER, ELLEN M.; RANDY L. DIEHL; and KEITH R. KLUENDER. 1986. Trading relations in speech and nonspeech. *Perception & Psychophysics* 39.129–42.
- PETERSON, GORDON E., and ILSE LEHISTE. 1960. Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* 32.693–703.
- PETURSSON, MAGNUS. 1976. Aspiration et activité glottale. Examen expérimental à partir de consonnes islandaises. *Phonetica* 33.169–98.
- PIERREHUMBERT, JANET B. 1980. The phonology and phonetics of English intonation. Cambridge, MA: MIT dissertation.
- . 1990. Phonological and phonetic representation. *Journal of Phonetics* 18.375–94.
- PORT, ROBERT, and JOHN DALBY. 1982. Consonant/vowel ratio as a cue for voicing in English. *Perception & Psychophysics* 32.141–52.
- REINHOLT PETERSEN, NEILS. 1978. The influence of aspiration on the fundamental frequency of the following vowel in Danish: Some preliminary observations. Copenhagen: Annual Report of the Institute of Phonetics, University of Copenhagen 12.91–112.
- . 1983. The effect of consonant type on fundamental frequency and larynx height in Danish. Copenhagen: Annual Report of the Institute of Phonetics, University of Copenhagen 17.55–86.
- REPP, BRUNO H. 1979. Relative amplitude of aspiration noise as a voicing cue for syllable-initial stop consonants. *Language and Speech* 22.173–89.
- RIORDAN, CAROL J. 1980. Larynx height during English stop consonants. *Journal of Phonetics* 3.353–60.
- RISCHEL, JØRGEN, and A. THAVISAK. 1984. A note on work in progress: Secondary articulation in Thai stops. Copenhagen: Annual Report of the Institute of Phonetics, University of Copenhagen 18.243–54.
- SAPIR, SHIMON. 1989. The intrinsic pitch of vowels: Theoretical, physiological, and clinical considerations. *Journal of Voice* 3.44–51.
- SAWASHIMA, MASAYUKI, and HAJIME HIROSE. 1983. Laryngeal gestures in speech production. The production of speech, ed. by Peter MacNeilage, 11–38. Berlin: Springer.
- ; ———; TATSUJIRO USHIJIMA; and SEIJI NIIMI. 1975. Laryngeal control in Japanese consonants, with special reference to those in utterance-initial position. Tokyo: University of Tokyo Research Institute in Logopedics and Phoniatrics, Annual Bulletin 9.21–26.
- SCHNEIDER, WALTER, and RICHARD M. SHIFFRIN. 1977. Controlled and automatic

- human information processing, I: Detection, search, and attention. *Psychological Review* 84.1-66.
- SHIFFRIN, RICHARD M., and WALTER SCHNEIDER. 1977. Controlled and automatic human information processing, II: Perceptual learning, automatic attending, and a general theory. *Psychological Review* 84.127-90.
- SILVA, DAVID J. 1991. A prosody-based investigation into the phonetics of Korean stop voicing. *Harvard studies in Korean linguistics*, IV, ed. by Susumu Kuno et al., 177-88. Cambridge, MA: Department of Linguistics, Harvard University.
- . 1992. The phonetics and phonology of stop lenition in Korean. Ithaca, NY: Cornell University dissertation.
- SILVERMAN, KIM E. A. 1986. F₀ segmental cues depend on intonation: The case of the rise after voiced stops. *Phonetica* 43.76-91.
- . 1987. The structure and processing of fundamental frequency contours. Cambridge: Cambridge University dissertation.
- SIMADA, ZYUN'ICI, and HAJIME HIROSE. 1971. Physiological correlates of Japanese accent patterns. Tokyo: University of Tokyo Research Institute in Logopedics and Phoniatrics, *Annual Bulletin* 5.41-49.
- SLIS, IMAN H. 1970. Articulatory measurements on voiced, voiceless and nasal consonants. *Phonetica* 21.193-210.
- , and ANTOINE COHEN. 1969a,b. On the complex regulating the voiced-voiceless distinction. *Language and Speech* 12.1:80-102, 11:137-55.
- SPROAT, RICHARD, and OSAMU FUJIMURA. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21.291-311.
- STEVENS, KENNETH N., and SHEILA E. BLUMSTEIN. 1981. The search for invariant acoustic correlates of phonetic features. *Perspectives on the study of speech*, ed. by Peter D. Eimas and Joanne L. Miller, 1-38. Hillsdale, NJ: Lawrence Erlbaum.
- STEVENS, KENNETH N., and SAMUEL JAY KEYSER. 1989. Primary features and their enhancement in consonants. *Lg.* 65.81-106.
- ; ———; and HARUKO KAWASAKI. 1986. Towards a phonetic and phonological theory of redundant features. Invariance and variability in speech processes, ed. by Joseph S. Perkell and Dennis H. Klatt, 426-63. Hillsdale, NJ: Lawrence Erlbaum.
- STEVENS, KENNETH N., and DENNIS H. KLATT. 1974. Role of formant transitions in the voiced-voiceless distinction for stops. *Journal of the Acoustical Society of America* 55.653-59.
- SUMMERFIELD, A. QUENTIN, and MARK P. HAGGARD. 1977. On the dissociation of spectral and temporal cues to the voicing distinction in initial stop consonants. *Journal of the Acoustical Society of America* 62.435-48.
- SUMMERS, W. VAN. 1987. Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses. *Journal of the Acoustical Society of America* 82.847-63.
- TULLER, BETTY, and J. A. SCOTT KELSO. 1990. Phase transitions in speech production and their perceptual consequences. *Attention and performance XIII: Motor representation and control*, ed. by M. Jeannerod, 429-52. Hillsdale, NJ: Lawrence Erlbaum.
- WEISMER, GARY, and DENISE CARISKI. 1984. On speakers' abilities to control speech mechanism output: Theoretical and clinical implications. *Speech and language: Advances in basic research and practice* 10, ed. by Norman J. Lass, 185-240. Orlando, FL: Academic Press.
- WESTBURY, JOHN R. 1983. Enlargement of the supraglottal cavity and its relationship to stop consonant voicing. *Journal of the Acoustical Society of America* 73.1322-36.
- , and PATRICIA A. KEATING. 1980. A model of stop consonant voicing and a theory of markedness. Paper presented at the 55th Meeting of the Linguistic Society of America, San Antonio.
- WOOD, SIDNEY. 1982. X-ray and model studies of vowel articulation. Lund: Lund Uni-

versity Department of General Linguistics, Phonetics Laboratory, Working Papers
23.

YOSHIOKA, HIROHIDE; ANDERS LOFQVIST; and HAJIME HIROSE. 1982. Laryngeal adjustments in Japanese voiceless sound production. *Journal of Phonetics* 10.1–10.

John Kingston
Linguistics Department
University of Massachusetts
Amherst, MA 01003
[kingston@cs.umass.edu]

[Received 5 May 1993;
revision received 2 November 1993;
accepted 5 March 1994.]

Randy L. Diehl
Department of Psychology
University of Texas
Austin, TX 78712
[diehl@psyvax.psy.utexas.edu]