Phonology **22** (2005) 395–436. © 2005 Cambridge University Press
doi:10.1017/S0952675705000667 Printed in the United Kingdom

*Laryngeal markedness and aspiration**

Bert Vaux
University of Wisconsin-Milwaukee

Bridget Samuels
Harvard University

We argue that the common phonological assumptions that (i) plain voiceless consonants are less marked than voiceless aspirates and (ii) the unmarked two-way stop system opposes unaspirated voiced and voiceless members are incorrect. A wide range of phonetic and internal and external phonological evidence suggests instead that (i) the maximally unmarked single-series stop is unspecified for laryngeal features and (ii) the unmarked two-way stop system opposes aspirated and unaspirated stops, and the aspirated series may be the unmarked member of this set.

1 Introduction

Phonologists generally assume that plain voiceless consonants (henceforth T) are less marked than voiceless aspirates (T^h) (Jakobson & Halle 1956, Greenberg 1966, Chomsky & Halle 1968, Maddieson 1984, Lombardi 1991, 1995, Barna 1998 and other works in Government Phonology, Silverman 1998, Burzio 2000, Iverson & Ahn 2001, etc.), and that the unmarked two-way stop system contrasts unaspirated voiced (D) and voiceless (T) members, as in Spanish. Systems containing a T^h series, as in English, Armenian and Turkish, are claimed to be marked in comparison (Maddieson 1984: 28; cf. Keating 1984, who proposes that all conventional two-series stop systems underlyingly oppose a voiced and a voiceless series). We suggest that these generalisations concerning markedness are incorrect. Following Rice (1999), Clements (2005) and Flemming (2005), there are many dimensions of markedness, which do not necessarily converge on a single target cross-linguistically in the way outlined above, and phonological patterning and feature inventories must

* An earlier version of this paper was presented by the first author as 'Systemic *vs.* feature-based markedness in laryngeal contrasts' at the LSA Annual Meeting in San Francisco in January 2002. Thank you to Fred Eckman, Morris Halle, Greg Iverson, Alec Marantz, Mickey Noonan, Keren Rice, Catherine Ringen and Joe Salmons for helpful comments.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

be considered carefully, since the typical criteria for evaluating markedness, such as implication and frequency, can be unrevealing. In light of the evidence this type of appraisal presents, following Steriade (1997: 22) and Wilson (2001), we propose that the maximally unmarked single-series stop is unspecified for laryngeal features, which is *not* the same as a voiceless unaspirated stop: the former may vary over voice onset time (VOT) and other phonetic parameters, whereas the latter is specified for particular laryngeal gestures. We argue moreover, based on evidence from acquisition, articulation, perception and first- and second-language phonology, that the unmarked two-way stop system opposes aspirated and unaspirated stops (T^h:T), and that the aspirated series (T^h) may be the unmarked member of this set.[1]

2 Traditional arguments for the unmarked status of plain voiced obstruents

Let us first consider the leading arguments that have been adduced in support of the claim that voiceless unaspirated obstruents (T) are maximally unmarked, as summarised in (1).

(1) *Traditional arguments for T being maximally unmarked for laryngeal features*
   a. Plain voiceless stops (T) are normally favoured in positions of neutralisation.
   b. They have restricted distribution in some languages.
   c. They are more common cross-linguistically than aspirates.
   d. They are more efficient aerodynamically and articulatorily, at least in initial position.
   e. Long-lag VOT (involved in the production of voiceless aspirates) is reduced in fast speech, whereas short-lag VOT (involved in voiceless non-aspirates) is unaffected.
   f. Children master and produce short-lag stops (T) before long-lag stops (T^h).

We propose that these arguments do not in fact support the claim that plain voiceless consonants (T) are less marked than voiceless aspirates

[1] Markedness in *three*-way series depends on whether the plain voiceless series is tense (as in Armenian) or not (Polish, Russian; Trubetzkoy 1958, Pisowicz 1976). In the former case the voiceless unaspirated series is marked, as it is in Lezgian and similar cases discussed in §3.4.1. We might actually except T^h to be least marked in a {D T T^h} system, because D requires muscle control and VOT synchronisation and T requires VOT control, whereas T^h requires neither of these. This issue is discussed in more detail in §3.2. Alderete *et al.* (1999) seem to imply that T would be marked in a three-way system, by extrapolation from their claim that 'in a three-tone system, the middle tone is most marked, dispersion-wise (see also Gnanadesikan 1997)'.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

(T^h) or the claim that the unmarked two-way stop system contrasts unaspirated voiced and voiceless members (D:T).

2.1 Neutralisation

Let us begin with the neutralisation argument in (1a). This is perhaps the argument most frequently invoked in support of the unmarkedness of plain voiceless obstruents; phonologists typically cite neutralisation to T in the syllable coda (Trubetzkoy 1958: 146–148, Greenberg 1966, Lombardi 1991, 1995, etc.), but sometimes include environments such as following /s/ in English (Harris 1994).[2] We shall see in this section, though, that for both phonetic and phonological reasons neutralisation is not always a strong argument for universal markedness.

Maddieson *et al*. (2001: 144) provide a typical example of this position in their analysis of Tlingit, where they state that the phonetically aspirated stops that occur in syllable codas must actually be in some sense *un*-aspirated, because 'for Tlingit to have an alternation between unaspirated consonants in onset position and aspirated consonants in coda position would be surprising, since in many of the languages that have aspiration contrasts in onset position, aspiration is disallowed in coda position (e.g., Thai, Korean, Cantonese, Navajo, Hupa, Tanana, and other Athabaskan languages)'. Their basic claim is that since Thai and the other languages mentioned neutralise to voiceless unaspirated stops in coda position, this must be the unmarked member of the opposition, and Tlingit must neutralise to this series as well, despite the phonetic evidence.

The neutralisation argument encounters three problems. First, segments occurring in positions of neutralisation are not necessarily voiceless unaspirated (T), the type argued to be the least marked universally. For instance, in Tlingit and many other languages detailed in §3.4.3, laryngeal neutralisation yields voiceless aspirates (T^h). Lezgian disallows plain voiceless stops in word-final position (and changes these to plain voiced stops in a class of monosyllabic nouns; Yu 2004). Similarly, several languages allow only voiceless stops after nasals (Hyman 2001), contrary to the commonly cited generalisation that plain voiced stops are preferred after nasals cross-linguistically (Pater 1999). Recently it has also been pointed out that any of the major places of articulation (labial, coronal, dorsal) can pattern as unmarked in some language, if one employs standard markedness diagnostics (Hume & Tserdanelis 2002, Hume 2003).

One reason that marked segments can surface in neutralisation contexts is failure for physical reasons to implement an articulation instruction outputted from the phonological component. For example, apropos of final devoicing in English, Tatham (1979) cites evidence from an electromyographic study by Tom Shipp of the San Francisco Veterans Affairs Medical Center in San Francisco showing that the larynx is actually in the

[2] Similar arguments have been adduced in support of plain voiceless stops being underlyingly unspecified for laryngeal features; cf. Clements (1985).

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

configuration for voicing in English word-final obstruents that ultimately surface as T, without vocal cord vibration. Morton & Tatham (1980) conclude from this that devoicing is actually aerodynamic, not phonological:

> The fact that in … American and British English … many, if not most, speakers tend to fail to produce vocal cord vibration throughout the phonetic realisation of the consonant … is irrelevant to phonology … .
>
> Cessation of vocal cord vibration in the realisation of a phonological [+voice] consonant in intervocalic environment is, of course, a non-programmed occurrence. By this we mean that it occurs as a non-intended artifact. Vocal cord tensing, raised subglottal air pressure, etc., are always present, and from this combination of phonetic features vocal cord vibration would normally occur spontaneously were it not for the fact that any oral occlusion will raise the supraglottal air pressure destroying the transglottal pressure drop which is conditional for vocal cord vibration.

Kim (1970) and Iverson & Salmons (1995) make similar observations for [spread glottis] in English, where the phonological instruction to spread the vocal folds may or may not result in greater airflow through the vocal tract, thus producing an eventual output of either T or T^h at a given point in time.

The second potential problem with the neutralisation argument is that neutralisation sometimes involves fortition, which is generally believed to involve an increase rather than a reduction in markedness. This has been proposed for German final 'devoicing', for example, by Iverson & Salmons (1999, 2003). We suggest in this paper that the voiceless aspirates produced by fortition of this type are not marked, though; for further discussion see §3.4.3.

A third problem with the neutralisation argument is that neutralisation is sometimes only partial or perceived rather than real (Labov *et al*. 1972, Port & O'Dell 1985, Slowiaczek & Dinnsen 1985, Smith 1997, Scobbie *et al*. 2000, Iverson & Salmons 2003). The misidentification of partial contrast preservation as neutralisation can be connected in part to the perceptual ambiguity resulting from absence of necessary phonetic cues (cf. Steriade 1997: 30, 2001). For this reason, we must treat all claimed cases of complete neutralisation with caution.

What the facts discussed in this section suggest is that one should not be so quick to invoke (perceived) neutralisation in support of a given segment being unmarked. The segment the linguist hears may not be the target output of the grammar – nor in some cases may it be the actual phonetic output – and hence is irrelevant for considerations of phonological markedness.

2.2 Distribution within phonological systems

Let us turn next to (1b), the argument that aspirates are marked because they have restricted distribution in some languages. In support of this

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

generalisation phonologists adduce facts such as the restriction of aspirates to word-initial position in the Indic languages South Mewari, Lamani and East Bengali (data from Masica 1991: 130), or the fact that only one aspirate per word is allowed in Harauti, South Mewari, some forms of Marwari and Konkani (Masica 1991: 130, 204), Grassmann's Law in Sanskrit and Proto-Indo-European (McCarthy 1988) and the Bantu language Makhuwa (Katupha's Law; Schadeberg 1999).

This type of argument is similarly inconclusive with respect to markedness, because both marked and unmarked members of a set may have restricted distributions of the sort just described. To take one example, Trubetzkoy (1958: 141) observes that in Tamil, obstruents surface as voiceless aspirated stops (T^h) initially, as voiced fricatives after vowels, as voiced unaspirated stops (D) after nasals and as voiceless unaspirated stops (T) after /r/. Though presumably some of these allophones are more marked than others, none can be said to have a less restricted distribution.

We also find cases where the putatively marked member of an opposition occurs in a less restricted set of environments than its unmarked counterpart(s). One such example is provided by Colin, a child with disordered English speech who controlled only a single place of articulation for stops; what he produced was not the coronal [t], which is widely believed to be the least marked stop, but rather the velar [k] (Bernhardt & Stemberger 1998: 590).[3]

2.3 Distribution across languages

Generalisation (1c), that voiceless unaspirated consonants (T) are more common cross-linguistically than voiceless aspirates (T^h), has gained additional credence in the literature since the publication of Maddieson

[3] Rice (1999) suggests that /k/ may actually be the least marked stop in such cases, citing the parallel of velars being targeted for assimilation in Chukchee (Odden 1987) and used for epenthesis in Uradhi (Crowley 1983) and Murut (Prentice 1971). (One might also add the merger of Proto-Lappic final stops to *-k* in Lapp; Mikola 1996.) One of the present authors has argued elsewhere (Vaux 2001) that consonant epenthesis is not a legitimate indicator of markedness, since epenthetic segments may arise from hypercorrection rather than emergence of the unmarked, but the Chukchee and Lapp cases are suggestive. Even if Rice's analysis of /k/ is correct, though, other examples illustrating the same point (namely that 'unmarked' segments can have more restricted distributions) can be produced, as noted by Hume & Tserdanelis (2002) and Hume (2003).

A further complication appears when we consider the case of Korean coda neutralisation, as Greg Iverson has pointed out. In this context Korean allows *only* [t] among the set of coronal obstruents in surface forms (in keeping with the predictions of standard markedness theory), but allows only coronals *other than* /t/ in this position in underlying forms (exactly opposite to the predictions of standard markedness). Interestingly, instances of coda [t] in new surface forms such as loanwords are assigned to the phoneme /s/, as in /intənes/ 'internet' (cf. Sohn 2001). What the Korean data suggest is that we minimally must distinguish between markedness in surface *vs*. underlying representations, and perhaps should abandon the distribution criterion altogether.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

(1984), which, based on a study of more than 400 languages, identified the generalisations in (2):

(2) *Cross-linguistic distribution of laryngeal systems* (data from Maddieson 1984: 28)
   a. One-series systems have T (plain voiceless) 98% of the time (49/50 languages).
   b. Two-series systems have T:D (plain voiceless *vs.* plain voiced) 72% of the time (117/162 languages).
   c. T shows up in 92% of languages, *vs.* 29% for T^h.

Rice (1999: 3) has pointed out, though, that 'standard diagnostics for markedness, implication and frequency, are, by and large, unrevealing' (we refer the reader to Rice's article for detailed justification and exemplification of this point). In the case of (2) the data are not only unrevealing but also inaccurate, as is the generalisation drawn from (2), which is employed for example by Lombardi (1991, 1995) in her defence of the unmarkedness of the plain voiceless series (T). Maddieson incorrectly identifies many languages with a voiceless aspirate series (T^h), some of which we have listed in (3), as having an unaspirated series (T) instead.

(3) *Some languages whose* T *series as identified by Maddieson* (*1984*) *is actually* T^h

Aleut (Taff *et al.* 2001: 239)	Manchu (Li 1996)
Amharic (Leslau 1995)	Tuvan (Li 1996)
Evenki (Li 1996)	Yakut (Krueger 1962)
German (Wagner 2002)	most likely several others
Goldi (Li 1996)	

This problem results from reading grammars, which often do not note aspiration, rather than listening to the languages (cf. Szigetvári 1996: 2).[4] Moreover, Maddieson does not take into account the observation by Harris (1996) and Dwyer (2000, 2001) that voice and aspiration systems show clear areal distributions; sampling by language family rather than by linguistic area skewed the results for this variable. (Jansen 2004: 82–83 discusses further problems with UPSID involving conflation and mislabelling of distinct laryngeal categories.)

One also wonders about generalisation (2a), according to which one-series systems almost invariably select the T series, voiceless unaspirated stops. (Maddieson 1984: 27 suggests moreover that the one language in his database that does not conform to this generalisation is incorrectly described.) Given what we know about dispersion and the active control of linguistic articulations (Tatham & Morton 1971, Liljencrants & Lindblom 1972; see also §3.3.1 below), we expect the single stop series in such

[4] This is unusual for Maddieson, whose work is normally based on collection of sophisticated and reliable phonetic data.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

languages to vary widely along VOT and other phonetic parameters; it would be odd for the stops to require the articulatory precision and effort necessary to stay within the 0–35 ms VOT range when phonological contrasts do not force this to be so. We predict based on the considerations presented in this paper (but contrary to the theory of 'the emergence of the unmarked'; McCarthy & Prince 1994) that the stops in these languages should vary between the VOT ranges of D, T and T^h. Further phonetic study will be required in order to determine whether this prediction is borne out, but preliminary data from native Australian languages discussed in Clements (2005) appear to provide confirmation.[5]

2.4 Ease of articulation

Argument (1d) asserts that plain voiceless stops (T) are most efficient from the aerodynamic and articulatory points of view (at least in initial position; Keating *et al*. 1983), and that short-lag stops (T) are less difficult to produce than voice-lead (D) or long-lag (T^h) stops (Kewley-Port & Preston 1974). Auzou *et al*. (2000: 137–138) provide the most detailed elaboration of this position:

> The articulatory gestures for short voicing lag stops are easier than for the other two types of VOT. Voicing lead requires more muscle gestures than those needed for short voicing lag stops. The production of initial stop consonants with voicing lead necessitates some mechanisms external to the larynx to sustain an adequate transglottal pressure drop during the closure. In contrast, the stop consonants with long voicing lag require more carefully controlled timing between the oral stop and laryngeal closure. For these consonants, the adduction of the vocal folds requires more complex muscle activity … . In producing the voicing contrast, a speaker needs to coordinate the timing of velopharyngeal closure, supraglottal articulator closure, vocal fold oscillation and supraglottal articulator release.

For these reasons, T is claimed to be less marked than T^h and D. It should be noted though that Kewley-Port & Preston (1974) only *hypothesise* that short-lag stops (T) are less difficult to produce, rather than having articulatory evidence for this claim. In fact, phoneticians have found that true voiceless unaspirated stops (T) actually require specific laryngeal instructions (Halle & Stevens 1971, Hutters 1985, Dixit 1987, Goldstein 2001). More specifically, Hutters (1985: 1–2) suggests that the production of voicelessness requires a particular type of glottal abduction controlled by the posterior cricoarytenoid and interarytenoid muscles (cf. also Lisker 1977, Westbury 1983). The complexity of voiceless unaspirated stops

[5] Disner (1983) provides a few examples of non-dispersed vowel systems, which appear to run counter to this prediction. As Vaux has argued elsewhere (2002b), though, in a neo-Darwinian theory of language sub-optimal systems of this type are predicted to be *possible*, but *less common cross-linguistically* than their more dispersed and hence more learnable counterparts.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

relative to other series may help explain the finding that English-speaking stutterers tend to produce voice-lead (D) rather than short-lag (T) stops (Viswanath & Rosenfield 2000).

Further evidence that plain voiceless stops (T) are actually more complex articulatorily than voiceless aspirates (T^h) appears when one considers the articulatory gestures involved in producing the two types of consonant. As we discuss in more detail in §3.2, short-lag stops (T) require precise regulation of VOT, whereas long-lag stops (T^h) do not; the speaker has great latitude in the degree of variation allowed in coordination of closure release and laryngeal control.

In sum, the weight of the articulatory evidence suggests not only that long-lag voiceless aspirates (T^h) are not *more* complex than plain, short-lag voiceless stops (T), but also that they are actually *less* complex.

2.5 Fast speech

Our next argument to consider is (1e), which is based on the observation that in fast speech long-lag VOT, as in T^h, is reduced, whereas the short lag of T is unaffected (Kessinger & Blumstein 1997, Koenig 2001: 1064). Under the assumption that markedness is reduced in fast speech registers (cf. Nathan *et al*. 1987, Kirchner 1994), we might infer that short lag is less marked than long lag. This chain of reasoning is problematic, though, because it is not entirely clear that markedness is invariably reduced in fast speech; in fact, it is well known that faster speech often produces *more* marked rather than less marked forms. A classic example is English nasal place assimilation (Clements 1985), which in faster speech registers produces complex alveo-labio-dental nasals in words like *infinite*, which in more careful registers would have a simple alveolar nasal. (See also the discussion of fast speech forms like [ptejɾow] for *potato* in Reiss 1997b.)

In other words, the behaviour of fast *vs*. careful speech says nothing one way or the other about markedness; in some cases fast speech appears to involve reduction of markedness, and in others it engenders greater markedness (presumably due to simplifications or enhancements along other cognitive or physiological dimensions). It is possible that markedness is simply orthogonal to rate of speech, and other factors are responsible for the apparent changes in markedness in these situations.

2.6 Child language

Moving on to child language (1f), a number of arguments have been adduced from this domain by Jakobson (1941) and subsequent researchers in support of the idea that T is maximally unmarked. Consider first the development of babbling in infants: aspirated stops (T^h) are claimed to be rare in babbling whereas plain voiceless stops (T) are common (Preston *et al*. 1967, MacNeilage 1980: 19). Irwin (1947: 398), for example, found that in the babbling of 11–12-month-old infants 'd' (apparently referring

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

to plain voiceless [t]) is the second most common sound produced, appearing as 20·58 % of all tokens, whereas 't' (i.e. voiceless aspirated [t^h]) is only fifth most common, at 4·34 %. The results appear to be comparable for slightly older deaf children: Carr (1953: 26) found in his study of the spontaneous vocalising of 48 five-year-old deaf children that the most common sound was 'b' (i.e. plain voiceless [p]), at 20·48 %, whereas 'p' (i.e. voiceless aspirated [p^h]) was only third most common, at 11·36 %.

Similarly, in early speech production the first stops produced by children (up to 24 months) are claimed by many researchers (Jakobson 1941, Burling 1959, Macken 1980, Macken & Barton 1980, Dinnsen 1992, Engstrand & Williams 1996) to be voiceless and unaspirated (T); Preston *et al.* (1967) found that long-lag stops began to occur more frequently at about age 2. Kewley-Port & Preston (1974) also found that mean VOTs for American children's production of /t/ remained shorter than adult values as late as 4;5 years. Some have argued as well that the target VOT range is controlled earlier for English-speaking children for 'voiced' stops (i.e. T) than for the aspirates (T^h) (Gilbert 1977, Allen & Hawkins 1980: 236).

In the domain of disordered child speech, researchers have observed that a deaspiration phenomenon is quite common; Compton (1976: 89), for example, found in a study of twenty children aged 5–6 with severe articulatory disorders that thirteen (65 %) displayed deaspiration (cf. also So & Dodd 1994, Bernhardt & Stemberger 1998: 584; Compton 1976: 72 mentions that all six of his normal child subjects showed deaspiration as well). In a similar study Dinnsen (1999: 657) observes that 'subject 29, age 4;11, illustrate[s] adultlike underlying representations with respect to the [voice] feature of initial obstruents [(4a)] but the failure to aspirate voiceless stops [(4b)]'.

(4)

	output	*target*	
a.	[bɪg]	[bɪg]	big
	[bæg]	[bæg]	bag
	[dɒg]	[dɒg]	dog
	[dʌk]	[dʌk]	duck
b.	[pɪg]	[p^hɪg]	pig
	[pɑɪ]	[p^hɑj]	pie
	[tut]	[t^huːθ]	tooth
	[tʌb]	[t^hʌb]	tub

Macken & Barton (1980) and Scobbie *et al.* (2000) note moreover that even after children acquire a two-way voicing contrast, both series initially fall within the short-lag range (at least for some children).

Under the Jakobsonian hypothesis that children's productions are unmarked relative to adult models, the facts just presented suggest (argument (1f) maintains) that short-lag stops (T) are less marked than long-lag stops (T^h). These facts suffer from two important drawbacks, however: some of the data are inaccurate or incomplete, and the interpretation of child performance (rather than competence) data is problematic.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

To begin with, the child-language data just discussed wrongly imply that children converge at an early stage precisely on short-lag stops (T). Smith (1973), Cruttenden (1985), Gierut & Dinnsen (1986) and others observe that it is common for children to actually first produce voiced stops (D) rather than plain voiceless stops (T) word-initially before vowels, and only later to begin producing T in this position. Eguchi & Hirsh (1969), Eilers *et al.* (1984), Tyler & Saxman (1991), Auzou *et al.* (2000) and others observe moreover that there is significant variation in child production of American English along the VOT dimension, randomly distributed from voicing lead (D) to long voicing lag (T^h) (Auzou *et al.* 2000: 140). Similarly, for children learning Spanish, what has been labelled as a plain voiceless stop (T) actually varies widely across both the short-lag (T) and long-lag (T^h) regions of VOT, with values from 0 to 140 ms (Williams 1977); the same was found in a recent study of children acquiring English, where only 69% of the tokens fell into the adult 30–105 ms range (Foulkes *et al.* 1999). Koenig (2001: 1059) notes that 'inspection of Kewley-Port & Preston's [1974] histograms shows that, although most of the children's /t/s fell in the short-lag range, *each child produced a few outliers with very long VOTs* (*up to 290 ms*)' [emphasis ours]. She adds that some children appeared to reach mature values by first 'overshooting' (producing very long positive VOT means) and then shortening back to adult values. Macken & Barton (1980) found longer VOT means in four-year-olds than adults and took this as further evidence for an overshoot phase in acquiring voiceless aspirates. Other authors as well have reported long means and occasional very long (>100 ms) positive values of VOT in children's productions of /p t k/ (Menyuk & Kiatt 1975, Gilbert 1977, Smith 1978). Eguchi & Hirsh (1969) attribute the significant variation in child VOT to the fact that the extremely precise temporal coordination involved in VOT requires a fully mature nervous system, an idea that will be developed in more detail when we discuss articulatory aspects of voicing and aspiration in §3.2.

The child-language facts just discussed are paralleled in normal adult speech. Williams (1980: 207), for example, found that English speakers range between lead lag (D) and short lag (T) in their production of the 'd' series of stops.

The same variability holds true for the disorder facts presented earlier; children who produce 'deaspiration' are typically not producing specially voiceless unaspirated stops (T), but rather stops that cover the VOT gamut from T to T^h. Compton (1976: 68–69) found in a study of Grace, a five-year-old with disordered speech, that in word-initial position she produced stops in the long-lag region 38% of the time, and stops in the short-lag region 62% of the time. In initial clusters she aspirated stops 54% of the time. We appear to be dealing in all of the cases just described not with voiceless unaspirated stops, but with stops unspecified for the relevant laryngeal features, with concomitant implementational variation along the lines set out by Steriade (1997, 2001), resulting from the fact that

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

the child has not yet mastered the relevant articulatory gestures (Ferguson & Farwell 1975, Goldstein 2001).

Disordered *adult* speech appears to manifest a similar phenomenon. Andrea Calabrese (personal communication) reports that the Italian-speaking aphasics he studied with Christina Romani produce the voiceless stop series, which in standard Italian are plain voiceless (T), as voiceless aspirates (T^h).

Returning to child speech, it has also been observed that there is variability across children for order of acquisition of laryngeal features and other acquisitional sequences.[6] Cruttenden (1985) suggests moreover that 'when /ptk/ develop, it is often the aspirated forms which occur first, because these phonemes are usually learnt first in word-initial position'. Goad & Brannen (2000) add that children initially aspirate word-*final* aspirates as well. In other words, it is inaccurate to say that children in English-speaking environments first acquire and/or produce short-lag stops (T); there is in fact evidence that such children initially master long-lag stops (T^h), at least in initial and final position.

Even if one *could* find evidence of the opposite, namely children who consistently produce short-lag stops for the aspirated series ('de-aspiration'), this would have to be confronted with the fact that *hyper*-aspiration, with VOT values of 100 ms and over, has also been claimed as a syndrome in both normal and disordered child language (Catts & Jensen 1983, Scobbie *et al*. 2000). In other words, we have child speech where the VOT is *shorter* than usual for the aspirate series (T^h), and child speech where it is *longer* than usual. Are we justified in building a markedness argument from just one of these two facts, to the exclusion of the other? According to Jakobson (1941), Smolensky (1996) and many others, we expect deviant children's forms to be less marked. Are both of these options (hyper-aspiration and deaspiration) then less marked than adult aspiration? The basic problem is that we know children may produce forms that are *less* marked than their adult targets, but we also know that they may produce *more* marked forms (Reiss 1997a). Given that the data cut both ways, attributions of markedness to child outputs ought to be interpreted with extreme caution.

We must also bear in mind that what should concern us as phonologists is the *target* of the articulation, not its implementation. This unfortunately is difficult to verify with children. What we really want to know for example is whether the muscle instruction for glottal spreading (the articulatory implementation of aspiration/long-lag VOT) is being sent in the children who are seemingly not aspirating. Kim (1970) shows that it is possible to have a spread glottis without significant airflow; we therefore

[6] Menn (1976) notes that the features to be acquired first may vary by place of articulation; at one point for example her child subject Jason had the output phonemic inventory /b t d k/, with a two-way contrast at the coronal place of articulation, a plain *voiceless* stop at the dorsal place of articulation and a plain *voiced* stop at the labial place of articulation.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

cannot be sure when we observe lack of airflow that the glottis is *not* spread.

The upshot of the child data considered in this section is that we must take great care in interpreting all such data, and when collecting child-language data we should place a premium on facts that reflect *competence* rather than *performance*. With these caveats in mind, the child-language facts do not in fact support (1f) – because (i) the data do not always support short-lag T being favoured, and (ii) it is not clear that children invariably diverge in the direction of lesser markedness – but rather point in favour of the theories advanced in this paper, because short-lag T is harder to implement and therefore less likely to be acquired successfully by the child.

2.7 Unwarranted phonologisation of VOT

Another problem with the studies invoked in support of the theory that plain voiceless stops (T) are maximally unmarked is that they rely on VOT as the element involved in distinguishing stop series (following an initial suggestion by Lisker & Abramson 1964). It is has been found, though, in many studies (Forrest & Rockman 1988; cf. also Hutters 1985: 17, Schiefer 1992) that VOT is not an overriding cue to stop voicing and aspiration; other cues include F1 values (Summerfield & Haggard 1977), formant transitions (Tsui & Valter Ciocca 2000), vocal cord vibration during closure (Stevens & Blumstein 1981), duration, tenseness, glottal spreading or constriction, intensity of release burst (Repp 1979) and articulatory force (Keating 1984, Avery 1997). Tsui (1996), for instance, found that aspiration noise was an important cue for both normal and severely hearing-impaired speakers of Chinese in the perception of aspiration, but VOT had no effect. We return to the role of VOT in perception in §3.3.2.

The theory that VOT is the culprit in voicing distinctions encounters severe phonological problems as well. Bhatia (1976) observes that the VOT theory predicts universal neutralisation of aspiration in word-final and preconsonantal position, since there is no onset of voicing in these situations. This prediction is incorrect; many languages, such as Hindi, maintain the contrast preconsonantally, and many (such as Standard Eastern Armenian; Vaux 1998b) maintain the contrast word-finally. The VOT theory also incorrectly predicts the universal absence of final aspiration processes; as we show in §3.4.3, such processes are in reality quite common (Bhatia 1976: 72–73 cites cases in Kashmiri and Amerind languages; Houlihan 1977: 236 adds Northern Ostyak (Rédei 1965) and Keresan (Spencer 1946)). Finally, Bhatia demonstrates on the basis of a battery of perceptual experiments that aspirates are in fact the most distinctive and most accurately identified stops in final position, contrary to the prediction of the VOT theory. We therefore expect aspirates to be more learnable than their unaspirated counterparts, a fact that plays an important role in the theory of markedness employed and developed in this paper.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

2.8 Interim summary

To summarise the arguments presented thus far, the claim that plain voiceless stops (T) are maximally unmarked rests on shaky ground. The data on which the claim is based are either incomplete (as in the case of child and disordered production data) or inaccurate (as with Maddieson's cross-linguistic survey of T *vs.* T^h), and the inferences drawn from them are not always justified (as in the cases of neutralisation and allophonic distribution, or the unwarranted phonologisation of VOT). Careful consideration of these factors suggests that the common conclusion about T being maximally unmarked is not as solid as generally presumed. On the contrary, there is sufficient evidence to support the adoption of the competing hypothesis that plain voiceless stops are marked relative to voiceless aspirates and unspecified stops.

3 Aspiration as unmarked

In §2 we argued that voiceless unaspirated stops, while claimed to be the most unmarked of the laryngeal series, do not pattern as such with respect to a number of factors. In this section we turn our attention to stops in opposition, arguing that (i) the unmarked two-way stop system actually opposes aspirated and unaspirated stops, and (ii) the aspirates may be the unmarked member of this set. This idea is consonant with the recent work of Iverson & Salmons (1995, 2003) and Avery & Idsardi (forthcoming), who suggest that a bare (i.e. unmarked) laryngeal node may be enhanced by [spread glottis], which apparently does not change the markedness relations in the system. A number of acquisition facts and articulatory, perceptual and phonological considerations support the two proposals advanced here.

3.1 Acquisition

While child-language data are sometimes misrepresented or misinterpreted, as we saw in §2.6, when considered in their entirety and in their proper context they can provide important insights into the relative complexity and difficulty of linguistic phenomena. Most relevant for our present purposes is the common and reasonable assumption that a less complex system will be learned faster and more accurately than a more complex system. With this in mind, it has been observed in the domain of first language acquisition that the two-way system acquired earliest by L1 learners is plain voiceless *vs.* voiceless aspirated ($T:T^h$) (Leopold 1947, Olmsted 1971, Cruttenden 1985, Goldstein 2001); children in an English-speaking environment acquire the $D/T:T^h$ contrast between 1;10 and 2;8 (Smith 1973, Macken & Barton 1980), whereas children learning Spanish do not master their D:T contrast until after the age of four (Macken &

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

Barton 1979).[7] Similarly, Deuchar & Clark (1996) found that a bilingual Spanish-English child acquired an adult-like voicing (i.e. D:T^h) contrast in English by 2;3, but had not yet done so for their Spanish D:T system.

Similar results, namely that the English-type system (D/T:T^h) is acquired faster and more successfully than a voiced/voiceless system (D:T), have been obtained for second-language learners in recent work by Fellbaum (1996) (for Spanish and English) and Khattab (2000) (for Arabic and English). This can be connected to Major & Kim's (1999) suggestion that learners have more difficulty acquiring similar sounds than dissimilar sounds, which they base on the finding that there is no difference between novice and experienced learners in their recognition and production of dissimilar sounds, which are acquired quickly, but there is a difference in the production and perception of similar sounds.

It is also worth noting that some dialects of languages whose voiceless stops are unaspirated have changed these into voiceless aspirates, even in the apparent absence of neighbouring languages displaying such systems.[8] Examples of this include the Basque Souletin variety of Basse-Soule, France (in at least some environments; Coyos 1994 and references therein), north-eastern Dutch dialects (Jansen 2004: 79), Cypriot Greek (Newton 1972), Cosenza Italian (Sorianello 1997), Salentino Italian (Andrea Calabrese, personal communication), Newscaster Malay (Poedjosoedarmo 1996) and perhaps Swahili (see discussion in Houlihan 1977). The fact that these languages have changed original D:T systems into D:T^h or T:T^h systems receives a straightforward explanation in the theory advanced in this paper, wherein voiceless aspirates (T^h) are less marked than plain voiceless stops (T), and aspiration contrasts (D/T:T^h) are easier to acquire and less marked than voicing contrasts (D:T).

3.2 Articulation

Turning next to the articulation of stops, based on simplistic considerations we might expect voiceless aspirates (T^h) to be less marked than D and T, because fully voiced stops (D) require specific muscle control and VOT synchronisation, and short-lag stops (T) require precise regulation of VOT, whereas long-lag stops (T^h) require neither of these types of precision. Closer inspection of the facts supports this idea.[9]

[7] Macken (1980) attributes this lag to the fact that the two series in Spanish contrast for the feature [continuant] as well as voicing, but this analysis remains controversial.

[8] This last clause is important; we want to exclude the possibility that the language has simply acquired its system from a linguistic neighbour by areal influence, as appears to be the case for the Dutch D:T system.

[9] The facts to be discussed here contravene Auzou *et al.*'s (2000: 137–138) claim that the articulatory gestures for short voicing lag stops (T) are easier than for T^h. They assert that 'stop consonants with long voicing lag require more carefully controlled timing between the oral stop and laryngeal closure. For these consonants, the adduction of the vocal folds requires more complex muscle activity'. As far as we can see, this assertion is directly contradicted by the facts and arguments presented here.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

Koenig (2001: 1059) observes that VOT measures the relative timing between laryngeal and supralaryngeal events (voicing/aspiration and closure/release respectively), and hence 'the sequence by which children attain adult-like values of VOT has been widely used as an indication of interarticulator timing control development'. The evidence suggests that the amount of interarticulator control required to implement a voiceless aspirate (T^h) is not as great as it is for a plain voiced (D) or voiceless stop (T). Goldstein (2001) observes that voiceless aspirates do not in fact require laryngeal–oral gesture coordination. It is specifically mastery of the aspiration *contrast* (T:T^h), not the segments on their own, that requires coordination of a laryngeal gesture (glottal opening–closing) with an oral constriction gesture. In general terms, an aspiration contrast requires coordination of only this single gesture pair, whereas a voicing contrast (D:T) requires more sophisticated coordination of multiple gestures (Kewley-Port & Preston 1974, Goldstein 2001). Auzou *et al.* (2000: 137–138) specify that in order to produce a voicing contrast, a speaker needs to coordinate the timing of velopharyngeal closure, supraglottal articulator closure, vocal fold oscillation and supraglottal articulator release. These facts help make sense of Gandour *et al.*'s (1986) observation that the voiceless aspirates are acquired before the voiced series in Thai and Hindi, and the fact that Parkinson's Disease and other speech disorders involving compromised articulatory control show significant increases in the amount and range of VOT in voiceless consonants (Özsancak *et al.* 1997 and references therein). Ageing, which also involves degeneration of articulatory control, produces a similar increase in VOT variability (Sweeting & Bacon 1982, Morris & Brown 1994). Tyler & Watterson (1991) speculate that the excess muscular tension exhibited by speakers with Parkinson's Disease inhibits appropriately timed approximation and initiation of vibration, leading to longer short-lag durations. To some extent the VOT modifications in Parkinson's might also result from a reduction in the coordination among the different structures in the speech production mechanism.

Morton & Tatham (1980: 7ff) observe along similar lines that the default state for CV coarticulation is to cause aspiration of voiceless consonants before vowels; a 'production instruction' is required to inhibit this universal phonetic rule and produce the closer alignment of release and voice onset required for short-lag stops in Italian, French and so on. Interestingly, they add (1980: 10) that inhibiting this universal phonetic rule is of limited utility, since the two-way contrast is easily maintainable without it, and therefore the production instruction is generally avoided in languages.

There is ample experimental evidence that voiceless non-aspirates (T) require active laryngeal control (Halle & Stevens 1971, Fischer-Jørgensen & Hirose 1974 (based on EMG data)), which makes sense, given the difficulty involved in coordinating release of closure with onset of voicing.[10]

[10] This notion finds a phonological parallel in Government Phonology, where it is assumed that voiceless stops are more complex/marked than voiced stops (e.g. Barna 1998: 15).

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

More precisely, Hutters (1985) concludes based on evidence from electromyography, photoelectric glottography and fibreoptic stills that the posterior cricoarytenoid and interarytenoid muscles are crucially involved in the production of plain voiceless stops (cf. Lisker 1977, Westbury 1983). Dixit (1987) and Steriade (1997) add that it is only *neutralised* (i.e. unspecified) obstruents that do not require complicated gestural coordination.

3.3 Perception

There are also many perceptual considerations that favour an aspiration contrast over one of voicing. In this section we consider enhancement, innate sensitivity to the long-lag aspiration boundary and perceptual salience.

3.3.1 *Dispersion/enhancement.* It has been suggested that certain systems favour dispersion or enhancement of the perceptual contrast between their members, notably by Lindblom (1986) for vowel systems. Enhancing a two-way contrast with aspiration may result in part from this phenomenon; increasing the acoustic–phonetic distance between voiced and voiceless phoneme categories has been claimed to accentuate their perceived contrast (Williams 1980: 201; cf. Avery 1997, Iverson & Salmons 2003). Similar notions are expressed by Keating's (1984) polarisation principle, according to which (*contra* Keating *et al.* 1983 and Holt *et al.* 2004: 1764) languages tend to maximise differences in VOT between contrastive series of obstruents, and by Iverson & Salmons (2003) and Avery & Idsardi (forthcoming). We in fact find many languages that oppose a fully voiced series to a voiceless aspirated series (D:T^{h}), such as Swedish (Ringen & Helgason 2002), some idiolects of English (Williams 1980, Singhaniyom 1999, Scobbie 2002),[11] Swahili (Polomé 1967) and many Western dialects of Armenian (Vaux 1998b). Since a D:T^{h} contrast is easier to perceive than a D:T contrast, we expect the former to be more easily learned than the latter, and hence more common over evolutionary time. We also expect to find cases where less-dispersed systems (D:T, T:T^{h}) evolve into more-dispersed systems. One such example is Modern Persian; Pisowicz (1987) observes that 'the emergence of aspiration [of the original plain voiceless series (T) in Persian] was probably caused by the tendency to increase the distance between the voiced and the voiceless members'.

The maximisation of contrast using aspiration can also be seen in rate of speech effects. Many studies have found that VOT values vary with speaking rate in normal speech production: VOT decreases as the rate

[11] Note, incidentally, that Lisker & Abramson (1964), which is one of the *loci classici* for the claim that plain voiceless stops (T) are unmarked, observe that one of their four English speakers has large voicing lead (i.e. D) for the 'voiced' series, and another does it some of the time. Their other two speakers use short lag (T).

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

increases (cf. Diehl *et al*. 1980, Miller 1981, Baum & Ryan 1993, Jäncke 1994). Slower speaking rates result not only in longer VOT but also in a wider range of VOT values (Miller & Volaitis 1989). Auzou *et al*. note that 'this effect is particularly prominent in English for voiceless stop consonants while values for voiced stops remain relatively stable. Thus, the difference between VOTs of voiced and voiceless consonants is reduced when the speaking rate increases' (2000: 141). The relevance of this empirical finding for our purposes is that when speakers have access to the requisite cognitive and articulatory resources, they prefer to disperse their stop series, particularly by increasing the amount of aspiration on the voiceless series.

Flemming (2005: 23ff) allows for enhancement of a voicing contrast via prenasalisation, implosion, etc., but does not mention the role played by aspiration, which, as we have just seen, is quite important in the production of stop contrasts. He also notes that enhancement of a voicing contrast should only happen when there is a contrast; when there is no contrast, in his scheme, the allophone of the relevant series that requires the least effort should surface, i.e. voiceless initially and voiced intervocalically and post-nasally (he cites Westbury & Keating 1986 in support of this position). We have seen already, and will see in more detail in the rest of this paper, that Flemming's version of dispersion theory falls short in this respect: other factors (such as perceptual enhancement and ease of gestural coordination) can trump voicing-related articulatory factors, even in the absence of contrast.

Returning to dispersion in systems where there *is* a contrast, though, it should come as no surprise that it is also difficult to implement a system that allows less room for dispersion. Importantly, in systems containing three or more series the *plain voiceless* series (T) is more constrained than the others in terms of VOT; the voice-lead series (D) can extend quite far in the lead direction and the long-lag series (T^h) can extend in the lag direction, but if the short-lag series (T) extends in either direction it will overlap with one of the other series. We schematise this dynamic in Fig. 1.

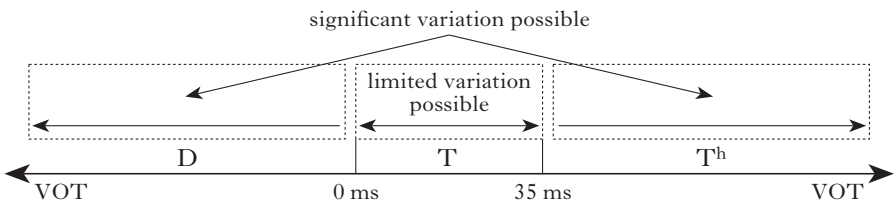


Figure 1
T constrained along the VOT axis relative to D and T^h.

Nepali, which has a standard Indic four-way contrast, provides a concrete demonstration of this dynamic at work:

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

	voice lead	short lag	long lag	voiced aspirate
mean (ms)	−76·96	14·63	70·55	−1·47
SD	15·33	7·87	16·37	25·87

Table I
VOT values in Nepali: means and standard deviations (from Poon & Mateer 1985: 42).

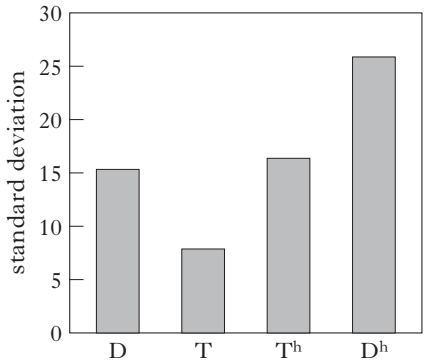


Figure 2
95% confidence intervals for VOT in the obstruent series of Nepali. Note the significantly smaller range of VOT variation for the short-lag (T) series.

As can be seen in Table I and in Fig. 2, the short-lag series shows a significantly smaller range of variation in VOT than do the other three series in Nepali, voice-lead (D), long-lag (T^h) and voiced aspirate (D^h). This is precisely what we expect on the basis of the VOT schematic in Fig. 1; the short-lag series has little room for variation, whereas the possibilities are much greater for the flanking series, D and T^h.

A similar situation occurs with the three stop series in Taiwanese. Hsu & Jun (1998) collected VOTs across three prosodic positions (intonational phrase-initial, word-initial and sentence-initial) in Taiwanese for aspirated /k^h/, unaspirated /t/ and voiced /b/. As shown in Fig. 3, they found that as the magnitude of the prosodic boundary increases, the VOT of /k^h/ increases and that of /b/ decreases, while the VOT for the middle stop series represented by unaspirated /t/ does not change. Based on these results, Hsu & Jun suggest that certain features associated with domain-initial segments are enhanced in order to enhance paradigmatic contrast among segments.

This phenomenon of short-lag stops (T) involving more precise control has an important consequence, observed by Davis (1995) for Hindi and other languages: whatever the languages being learned, children produce contrasts in minimal pairs with *large* VOT differences (such as English) *before* producing contrasts in minimal pairs with *small* VOT differences

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

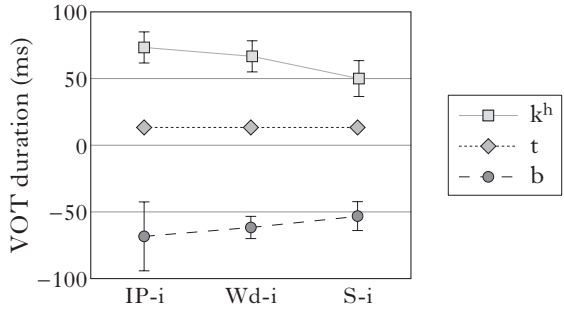


Figure 3
VOT duration for Taiwanese stop categories in three prosodic positions (Intonational Phrase-initial, Word-initial and Sentence-initial) (adapted from Hsu & Jun 1998: 82).

(such as Spanish), which require more precise gestural coordination. This dovetails well with the perceptual facts, to which we turn in the next section.

3.3.2 *Perceptual salience.* Our next perceptual argument begins with the oft-cited fact that *all* linguistic groups have their major discriminiation peak at the English (i.e. long-lag) VOT boundary for both children and adults. Beach *et al.* (2001) found for instance that monolingual speakers of Australian English and bilinguals who also spoke Greek were able to successfully distinguish Thai [b]:[pʰ] and (crucially) [p]:[pʰ], but not [b]:[p]. Eimas *et al.* (1971) obtained similar results for one- and four-month-old infants, who in their study discriminated [pa] and [pʰa] categorically, but not [ba] *vs.* [pa]. Similarly, Lasky *et al.* (1975) found that Guatemalan six-month-olds could discriminate voicing distinctions, but their perceptual category boundaries were closer to those of English-learning infants (i.e. at the long-lag VOT boundary, ca. 35 ms) than to those of their Spanish-speaking parents (at the short-lag boundary). Adult speakers of relevant languages, such as Spanish and Kikuyu, have a smaller perceptual peak at the Spanish border in addition to a main peak at the long-lag boundary (Lasky *et al.* 1975, Streeter 1976, Eilers 1980: 27–28, Willaims 1980: 208), but all of the above findings clearly agree on the English (long-lag) boundary being the most perceptually salient for both L1 and L2 learners.

This asymmetry in acquisition time has been connected by some to the well-known fact that the long-lag VOT boundary also exists in chinchillas (Kuhl & Miller 1975, 1978), rhesus macaques (Waters & Wilson 1976, Kuhl & Padden 1982), budgerigars (Dooling *et al.* 1989) and Japanese quail (Kluender *et al.* 1987), suggesting that humans (and perhaps all animals) are innately predisposed to categorise acoustic stimuli relative to this benchmark and thus receive a head start in acquiring an aspiration system (or, put differently, are more likely to acquire it successfully).

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

Regardless of whether or not the animal findings are relevant to the human case, the long-lag boundary appears to be an especially salient boundary in an *acoustic* sense (Stevens & Klatt 1974, Eilers & Minifie 1975, Williams 1980: 210–211, Aslin *et al*. 1983, Deuchar & Clark 1996, Vihman 1996: 74ff). Many phonetic features that have been found to cue a voicing/aspiration contrast are available at the long-lag *vs*. short-lag boundary but not at the voice-lead *vs*. short-lag boundary, including:

(i) formant cutback (presence/absence of transition information for the first formant) (Liberman *et al*. 1958, Lisker 1975, Tsui 1996);

(ii) spectral characteristics of the release burst, including aspiration noise between release and the onset of periodic vibration (Winitz *et al*. 1975, Williams 1977, Hutters 1985: 17);

(iii) differences in the degree and temporal extent of formant transitions (Cooper *et al*. 1952, Stevens & Klatt 1974, Summerfield & Haggard 1974);

(iv) differences in F0 change following release (Haggard *et al*. 1970, Fujimura 1971).

By contrast, voicing-lead cues are characterised by lower energy and are less perceptually salient (Williams 1980: 210).

Silverman (1995) notes moreover that 'the optimal timing relation between contrastive gestures is that which maximizes auditory nerve response, resulting in a more salient percept. As a sudden onset of spectral energy produces a heightened auditory nerve response, those articulatory configurations which produce an abrupt rise in spectral energy are most valued by phonological systems.' The fact that voiceless aspirates are more valued in this sense finds support in the perceptual study of Whalen *et al*. (1997), where 'in tests with non-words, aspirated [p^{h}] was consistently preferred over plain [p] without regard to which syllable was stressed. Additionally, the unaspirated stop was somewhat more difficult to discriminate from other allophones in a categorial AXB task than the aspirate was.' (The same tests with actual words showed preference for the proper allophone.)

The perceptual facts discussed thus far also dovetail nicely with Stevens' Quantal Theory (1972, 1989). The heart of Stevens' theory is the idea that linguistic contrasts tend to target 'quantal' acoustic regions, where changes in articulation have little effect on the speech output. Linking a linguistic category to a quantal region allows for a greater margin of articulatory error in implementing the category, and thereby enables the category to remain relatively stable over time. In the terms of the neo-Darwinian model employed in this paper, linguistic sound categories can be either quantal or not, but for the aforementioned reasons quantal categories are more likely to appear and to survive over evolutionary time than their non-quantal competitors. Applying this model to the case of short-lag (T) *vs*. long-lag (T^{h}) stops, the T category allows very little room for production errors (cf. the discussion of Fig. 1 and Table I above), whereas the T^{h} category allows a great deal of articulatory and acoustic wiggle room (via coarticulation, gestural mis-timing, speech errors, etc.) without compromising successful conveyance of category membership.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

What we conclude from all of this evidence is that there appears to be an innate sensitivity to the long-lag VOT boundary, which is supported by a host of perceptual cues that are not shared by the short-lag VOT boundary. We suggest that these two perceptual considerations reinforce the unmarkedness and/or greater learnability of the aspiration opposition. Two findings from studies of animal auditory perception may support this argument. Holt *et al*. (2004: 1763) observe that 'behavioral experiments with infants, adults, and nonhuman animals converge with neurophysiological findings to suggest that there is a discontinuity in auditory processing of stimulus components differing in onset time by about 20 ms', and suggest that perceptual boundaries close to this temporal value are easier to learn. Phillips (2001: 718) buttresses this argument with a detailed exposition that is worth quoting at length:

> A specific version of this hypothesis has been advanced by Steinschneider, Sinex and their colleagues … . According to this hypothesis, category boundaries for voicing are a natural consequence of the way in which specific groups of auditory neurons respond to events occurring in rapid succession. It is well-known that sensory neurons are often 'tuned' to specific stimulus features, to which they respond selectively. However, the act of responding to a preferred stimulus typically has the effect of *desensitizing* the neurons for a short period (the *inactivation period*), during which time the neurons cannot respond to new instances of the preferred stimulus. Steinschneider and Sinex suggest that the category boundary for VOT corresponds to the inactivation period for certain groups of auditory neurons which respond to both the noise-burst and the voicing onset. At VOTs above around 30 ms, the relevant cells can respond independently to the release and to the voicing onset, yielding a 'double on' response pattern. At shorter VOT values, on the other hand, the same cells respond only to the consonantal release, and cannot respond independently to the onset of voicing, yielding a 'single on' response pattern … . Support for this idea comes from direct recordings in primary auditory cortex of macaques … and humans … and the auditory nerve of chinchillas … . The 'double on' account of voicing suggests that voicing categories are based on low-level properties of the auditory system.

In sum, both the neural evidence adduced by Phillips and the acoustic and perceptual evidence that we have presented support the position that long-lag stops (T^h) are more perceptually salient and easy to identify than short-lag stops (T) are.

3.4 Phonological behaviour

The acquisition and phonetic facts considered thus far support our aspiration theory, but, as Rice (1999) has stated, the most important type of evidence in developing an analysis of markedness should come from phonological patterning. In the remainder of this paper we show that

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

the phonological evidence strongly supports the aspiration hypothesis as well.

3.4.1 *Previous proposals for voiceless aspirates being unmarked.* The idea that aspiration can be phonologically unmarked is not entirely new. Concrete proposals of this sort have been made by Trubetzkoy (1958: 263) for Lezgian, Pisowicz (1976: 88–89) for Eastern Armenian, Howe (2000: 31–32) for Oowekyala, Bach (1991) for Haisla, and Avery (1997) and Avery & Idsardi (forthcoming) for English.

To take one concrete example, Oowekyala, like Lezgian and Haisla, has a D:T^{h}:T^{ʔ} system; voiced stops and affricates occur only before tautosyllabic sonorants, and voiceless aspirated stops are unrestricted in their distribution (Howe 2000: 32). Based on the elsewhere status of the voiceless aspirates and their appearance in neutralisation contexts, Howe states that 'there is compelling evidence that the "aspirated" stops of Oowekyala are phonologically unmarked, while the phonetically "plain" stops are phonologically voiced … under this view of laryngeal contrasts, it is claimed that aspiration is no more than a phonetic property of stops and affricates with unmarked laryngeal specification in Oowekyala. Conversely, it is claimed that the lenis pronunciation of unaspirated stops … reflects the fact that these segments are phonologically marked ([voice] or [constricted])' (Howe 2000: 31–32).

In addition to Oowekyala, several other indigenous languages of the Pacific Northwest employ aspiration as the elsewhere case for stops. Czaykowska-Higgins & Willett (1997) report that Nxaʔamxcín stops are unaspirated in onsets; elsewhere they surface as aspirated. Cho & King (2003: 203) state that Bella Coola has the same distribution.

3.4.2 *Aspiration as default in English.* English may also be a system of this type, depending on how one treats the aspirated stops. Traditional analyses of English stop allophony of the sort outlined in (5) maintain that the aspirates are underlyingly plain, and receive aspiration by a regular aspiration process.

(5) /T/ → [T^{h}] foot-initially; [T] is the elsewhere case (cf. Kiparsky 1979, Clark & Yallop 1995, etc.)

Spencer (1996: 210) and others have pointed out, though, that apart from when they immediately follow /s/, any voiceless stop can optionally be aspirated.[12] Spencer therefore suggests that aspiration is actually the elsewhere case with English stops, showing up for example when the optional

[12] Cf. also Orton & Dieth (1962–1971), one of the first sources to systematically notate final aspiration in spoken English varieties. Kreidler (1989), Gee (1993: 127) and Davenport & Hannahs (1998) state that aspiration is optional in word-final position. Local (2002) states that (at least in British English) aspiration targets voiceless stops specifically in (conversational) turn-final position. In a nine-minute conversational sample he found 206 words ending in voiceless plosives, of which 61 occurred in turn-final position and 145 in turn-medial position. Of the turn-final tokens only one exhibited no aspiration; of the turn-medial tokens only four (3%)

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

rules of flapping and glottalisation do not apply.[13] In this analysis the rule responsible for the distribution of aspiration in English can be represented as in (6):

(6) $\emptyset \rightarrow$ [+spread glottis] / elsewhere (Spencer 1996: 211)

(The actual facts of aspiration in English are more complicated than this, but do not interfere with the basic argument being advanced here; the interested reader should consult Iverson & Salmons 1995 and Vaux 1998a, 2002a.)

According to McCarthy & Prince (1994) and McCarthy (2000) among others, segments selected for the elsewhere case – as aspirates are in (6) – are by hypothesis unmarked:

> Classic OT grammars share a general property called *harmonic ascent* … . A classic OT grammar … is a ranking of markedness and faithfulness constraints, and nothing else. Because violation is minimal, unfaithfulness is only possible to achieve markedness improvement relative to some language-particular ranking of the markedness constraints in UG. So, if a language has an unfaithful mapping /A/ $\rightarrow$ B, then B must be less marked, relative to that language's hierarchy, than the fully faithful candidate A (McCarthy 2000: 516).

If McCarthy is right, this entails that aspirated stops (T^h), since they are the elsewhere case, are phonologically unmarked relative to plain voiceless stops (T) in English.

One could say on the other hand that aspiration is underlying in English and lack of aspiration is derived, as suggested for example by Houlihan (1977: 236) (though she appears to ultimately reject this analysis), Iverson & Salmons (1995) and (in a slightly different form) Hale & Reiss (1998); one would then get the unmarked unaspirated form showing up in neutralising positions (cf. the analysis of Ossetic in Odden 2001: 52). Almost all analyses, however, assume that aspiration is not underlying in English, in accordance with the time-honoured assumption that predictable allophonic information is not stored in lexical entries. We follow this assumption here,[14] which leaves us with the conclusion (supported by the rest of the arguments adduced in this paper) that aspiration is unmarked.

---

exhibited aspiration, and three of these were followed by words beginning with aspiration (just heard, about half, cannot hear).

[13] In order to say that aspiration is the elsewhere case, one has to defend the existence of word-final aspiration (as some do; cf. Wells 1982: 46, Kreidler 1989, Vaux 1998a, Coleman 2003), which Laver (1994) and Odden (1996) claim is not actually aspiration. Greg Iverson (personal communication) supports Laver's idea, saying that final aspiration is impossible, since aspiration is measured in terms of VOT and there is no onset of voicing after word-final stops. The problem with this idea of course is that aspiration in the *phonological* sense is not actually VOT, but the result of spreading the vocal folds (Halle & Stevens 1971); VOT is but one of several possible *acoustic* manifestations of vocal fold spreading, as we saw earlier.

[14] Ignoring for the moment the fact that non-aspiration is also predictable in English (Lisker 1984). See Vaux (2002a) for further discussion.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

3.4.3 *Aspiration as product of neutralisation.* Our last phonological argument comes from neutralisation, which produces voiceless aspirates in a wide range of languages, some of which are given in (7). If we assume, as do supporters of 'the emergence of the unmarked' (McCarthy & Prince 1994), that the unmarked member of the opposition surface in positions of neutralisation, then we must conclude that voiceless aspirates (T^h) are unmarked relative to the other stop series in languages of the sort to be discussed in this section. (This conclusion must of course be tempered with the cautions regarding neutralisation that we raised in §2.1.)

(7) *A sampling of languages in which voiceless aspirates are produced by neutralisation*

a. *German*
Final stops, regardless of underlying voicing specification, are voiceless and (optionally) aspirated (Alber 2001, Wagner 2002).

b. *Turkish*
Final stops, regardless of underlying voicing specification, are voiceless and aspirated (Kopkalli 1993, Jannedy 1995, Gordon 1998).

c. *Tarascan*
Obstruents aspirate in neutralising position, e.g. post-obstruent (Foster 1969: 30–33, Steriade 1997: 67).

d. *Pit River*
Aspiration appears in the neutralisation case (syllable coda and preconsonantal; Nevin 1998).

e. *Danish*
[T^h] and [D̥] vary freely, [T^h] being used mainly in absolute final position (Fischer-Jørgensen 1980: 208).

f. *Swedish*
Stops aspirate in final/pre-pausal position (Johansson *et al.* 2001).

g. *Veron*
Final voiceless stops are neutralised as aspirated (Grierson 1906: 17).

h. *Yucatec Mayan*
Syllable-final obstruents aspirate (Straight 1976, Blevins 1995).

i. *Tojolabal*
Voiceless stops optionally aspirate in phrase-final position (Furbee-Losee 1976: 193).

j. *Yapese*
Voiceless stops aspirate word-finally; elsewhere they are unaspirated (Jensen 1977: 27, 38).

k. *Koyukon*
Has phonemic voiceless unaspirated (T), voiceless aspirated (T^h) and ejective ($T^ʔ$) stop series. 'There is no contrast between Koyukon plain and aspirated stops syllable-finally and the aspirated symbol is used in the writing system' (Marlow 2000: lxviii).

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

l. *Hupa*
Has phonemic voiceless unaspirated (T), voiceless aspirated (T^h) and ejective ($T^{\text{ʔ}}$) stop series. The T and T^h series surface as slightly aspirated in final position (Golla 1970: 26).

m. *Ahtna*
There is no contrast between aspirated stops and plain stops in word-final position: 'the word-final plain/aspirated stops sound about like word-final voiceless stops, when released, in English words such as *bet*, *pick*' (Kari 1990: 13–14).

n. *Quiegolani Zapotec*
'In utterance-final position the distinction between [voiced and voiceless obstruents] is generally lost, and both lenis and fortis members occur unvoiced with aspiration on stops and affricates' (Regnier 1993: 44).

o. *Icelandic*
Northern Icelandic dialects neutralise T and T^h series to T^h word-medially (Sidney Allen, personal communication).

p. *Tlingit*
Utterance-final non-ejective stops are aspirated (Maddieson *et al.* 2001: 143–144; note that final non-ejective stops in Tlingit orthography are written <t k>, etc. (i.e. aspirated), not <d g>, etc. (i.e. unaspirated)).

q. *Klamath*
See following discussion.

r. *Sierra Popoluca*
See following discussion.

s. *Kashmiri*
All plain voiceless stops are replaced with aspirates word-finally (Grierson 1906, Syeed 1978, Masica 1991: 204; see below).

Let us consider the last three of these in more detail. Beginning with Kashmiri, Grierson (1906: 17) observes that 'when a word in Veron or Kashmiri ends in an unaspirated surd [i.e. voiceless] consonant, the latter is always aspirated'. Syeed (1978) argues that this generalisation results from a neutralising rule of word-final aspiration, whose effects can be seen in (8). (We have added the underlying representations in (8) and (9).)

(8) *Kashmiri aspiration* (Syeed 1978: 535)

	UR	*nom*	*dat*	*agentive masc*	
a.	rat	rat^h	rat-as	rat-an	'blood'
	kaʈ	$kaʈ^h$	kaʈ-as	kaʈ-an	'sheep'
	haːk	$haːk^h$	haːk-as	haːk-an	'greens'
b.	nak^h	nak^h	nak^h-as	nak^h-an	'shoulder'

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

The dative and agentive forms in (8) show an aspiration contrast in the final root consonant that is neutralised to a voiceless aspirate when this consonant is placed in word-final position, as can be seen in the nominative. Syeed demonstrates that these facts require a final aspiration rule; attempting instead to derive the unaspirated forms by means of a de-aspiration rule would fail to capture insightfully the contrastive behaviour of minimal pairs of the sort in (9).

(9) *Kashmiri minimal pairs* (Syeed 1978: 535)

	UR	*nom sg*	*dat pl*	*agentive pl*	
a.	wat	wat^{h}	wat-an	wat-au	'way'
b.	kat^{h}	kat^{h}	kat^{h}-an	kat^{h}-au	'story'

The synchronic system closely reflects the historical development of the two series; cf. Kashmiri /kat^{h}/ < Old Indic /kat^{h}-ā/ 'story', Kashmiri /wat/ 'way' < /vartman/.

We can informally render the synchronic aspiration rule responsible for the alternations in (8) and (9) as (10):

(10) [+cons, −son, −voice, −cont] → [+spr gl] / __ #

This aspiration rule acts on loanwords as well: cf. Urdu /zaːt/ (a surname) → Kashmiri /zaːt^{h}/ (Syeed 1978: 539), Urdu-Persian /muhabbat/ 'love', /mehnat/ 'labour' → Kashmiri /mahabbat^{h}/, /mehnat^{h}/ (Syeed 1978: 540). Syeed adds that underlying and derived aspiration have the same acoustic manifestation in Kashmiri, and that 'the aspirated consonants in various positions [i.e. initial, medial, and final] were perceptually identified [by native speakers] as similar' (1978: 540). This latter fact suggests that the aspirates form a single psycho-phonological category for native speakers of Kashmiri, much like the orthographic facts suggested for Tlingit in (7p) above.

A relatively well-known case of final aspiration is Klamath, exemplified in (11), where, according to Blevins (1993: 246), 'word-finally, neutralization of obstruents is to the aspirated series not the plain voiceless series'. (Klamath, like Oowekyala, has a three-way laryngeal contrast T:T':T^{h}.)

(11) *Klamath word-final neutralisation to* T^{h}

/n'ep^{h}/	'hand'	→	[n'ep^{h}]	(cf. /n'ep^{h}eːʔa/ 'puts on a glove')
/ntʃ'ek'/	'in little bits'	→	[ntʃ'ek^{h}]	(cf. /ntʃ'ek'aːni/ 'small, little')
/nkak/	'turtle sp.'	→	[nkak^{h}]	(cf. /nkakam/ 'turtle's')

The same system appears in Sierra Popoluca, where stops are in coda position are neutralised to the voiceless aspirate (T^{h}) member of the series, as exemplified in (12) (cf. Elson 1947, Kenstowicz & Kisseberth 1979: 42–43).

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

(12) *Sierra Popoluca coda aspiration*

a. pet^{h}kuy 'broom'
pet^{h}paʔ 'he sweeps'
pet:a:p^{h}' 'it is being swept'

b. kek^{h}paʔ 'it flies'
nɨk^{h}paʔ 'he goes'
hu:t^{jh} 'where'
mak^{h}tiʔ 'ghost'
mok^{h} 'corn'

c. t^{j}u:kiʔ 'turtle'
ikapun 'his barrow'
kuy 'wood'

The facts of neutralisation in adult grammar that we just presented are paralleled in child grammars. Goad & Kang (2003: 125) cite the following examples (we omit their third example, from German, because *adult* German also aspirates final voiceless stops):

(13) *Aspiration in first language acquisition*

a. *English: Hildegard* (Leopold 1939: 108)
Word-final stops are 'strongly aspirated'
Target language: final stops may be unreleased, released and unaspirated, or aspirated
[be-t^{h}] 'Bates' (1;7)
[mit^{h}] 'meat' (1;10)
[bok^{h}] 'broke' (1;10)

b. *Québec French: Théo* (Rose 2000)
Target language: final consonants are not overtly released, unlike in European French

output	*target*		
[pɪc^{h}]	[pɪk]	'(it) pricks'	(2;05.11)
[əʃɛt^{h}]	[fʊʁʃɛt]	'fork'	(2;05.29)
[b̥ɪsɪk^{h}]	[bɪsɪk]	'bicycle'	(2;06.12)

We find the same pattern in disordered child acquisition. Compton's (1976) survey of 20 disordered children aged five–six revealed not only the familiar fact that 40% of them showed final devoicing, but also that 20% of them showed final *aspiration* (1976: 91). As we mentioned earlier, we must treat child production data with caution because they may reflect performance effects rather than linguistic competence, but the fact that in this case the normal and disordered child speech data mirror the evidence from normal adult phonological systems is suggestive.[15]

[15] Given the theory we have been advocating in this paper thus far, what we should really expect in positions of neutralisation is underspecification for aspiration, with the relevant segments then free to vary in degree of aspiration (cf. Wilson 2001 on voicing neutralisation, and Clements 2005 on voicing in Australian languages). To

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

As stated at the outset of this section, if we assume that the unmarked member of the opposition surfaces in positions of neutralisation (bearing in mind the cautions concerning neutralisation raised in §2.1), we must conclude that voiceless aspirates are unmarked relative to the other stop series in languages of the sort in (7)–(13).

3.4.4 *Second-language facts.* One finds the same phenomenon of neutralisation to voiceless aspirates in word-final position in second-language production. Stampe (1978), for example, makes the intriguing observation that languages without final and/or coda consonants devoice when they borrow such consonants (cf. also Eckman 1981, Altenberg & Vago 1983 (Hungarian–English), Flege *et al.* 1987, Weinberger 1987, Edge 1991, Yavaş 1994, Broselow *et al.* 1998 (Chinese–English)); we would add to this that at least in the case of several Japanese colleagues and students, stops in this position are not only voiceless but also aspirated.[16] The papers

---

take a concrete example, Steriade (1997: 30) states that 'we may predict that the neutralization of such distinctions [i.e. post-aspirated *vs.* unaspirated] will *always* occur in unreleased word-final and pre-obstruent stops [since release cues are essential to the perception of post-aspiration]'. Steriade is careful to avoid specifying what she means by neutralisation of distinctions here, though she implies that the product of neutralisation in this case would be an unaspirated segment, since children acquiring the language would not hear the cues to aspiration and therefore would not postulate its existence in their phonological system. (The same prediction is made by standard OT by virtue of 'the emergence of the unmarked' (McCarthy & Prince 1994), which requires that the unmarked pole of an opposition emerge in positions of neutralisation.) Given the nature of her argument, Steriade might want to say that all languages that neutralise to aspirates (such as Klamath) require release, whereas all those that neutralise to non-aspirates (such as Sanskrit) require *un*release (cf. Pierrehumbert 2000: 14: 'languages such as Urdu which have a full set of stops in all positions often require stop releases whether or not a vowel follows'). Cases like English would not be covered by this generalisation, since release is optional in many cases and mandatory in others; we suspect Steriade would predict that in optional cases such as this the neutralisation could go in either direction.

We suggest, though, that the correct interpretation of Steriade's generalisation should focus on the *absence of contrast*, rather than specifying aspirates or non-aspirates as the output of neutralisation. In this view any individual member of a set can appear in a position of neutralisation, as long as there is no contrast. This is essentially the situation we find in Rotokas (Firchow & Firchow 1969) and the Athabaskan language Tanacross: 'phonologically voiceless codas are phonetically voiceless and aspiration is not contrastive. Thus [ɬɛt], [ɬɛtʰ] and [ɬɛtˀ] are all acceptable pronunciations of /ɬɛt/ 'smoke'' (Holton 2000: 25).

We must also bear in mind that Steriade's cue-driven theory of neutralisation cannot be applied successfully to all cases – for example, (i) Ossetic neutralises (and deletes) aspiration on either side of an adjacent obstruent (Odden 2001: 52), (ii) New Julfa and several other dialects of Armenian neutralise aspiration after obstruents word-finally, which is a position where release is mandatory, as in English (Vaux 1998a), and (iii) Bangla neutralises aspiration but not voicing in final position (Vijayakrishnan 1999), even though the release cues that putatively drive voicing neutralisation should also require the neutralisation of aspiration. This therefore may not be the correct answer for any case of neutralisation.

[16] Cf. also Ross (1994). An anonymous reviewer attributes this phenomenon to there being a devoiced epenthetic vowel following, but the quality of this aspiration differs from the epenthetic vowel, and moreover if there were an epenthetic vowel (i) it

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

that notice final devoicing typically do not indicate whether these devoiced consonants are also aspirated (cf. e.g. Broselow *et al*. 1998), but aspiration has in fact been noted in the following cases:

(14) *L2 neutralisation to voiceless aspirates in word-final position*
  a. Mandarin speakers' acquisition of English (Heyer 1986) and French (Steele 2002).
  b. Korean speakers' acquisition of English (Pyun 1999).
  c. Icelandic speakers' acquisition of English (Mulford & Hecht 1980).
  d. Italian speakers' acquisition of English (Prator & Robinett 1972: 89).

The significance of the cases in (14) involving English acquisition is clouded by the aforementioned fact that English itself has optional final aspiration; the aspiration that shows up in learners' speech may therefore result from successful acquisition of the target rule rather than an interlanguage rule of aspirating neutralisation. Steele's case with Mandarin learners of French (16a) is more clear: since neither French nor Mandarin has a rule of final aspiration, this rule must have been postulated by the learners (or be a reversion to their UG system, depending on one's theory).

3.5 Language change

Evidence for change of plain voiceless stops in favour of voiceless aspirates can also be found in the historical record. Cypriot Greek, for instance, develops aspiration of original voiceless stops in absolute initial position, word-initially when preceding word ends in a vowel, and word-medially after a vowel (Houlihan 1977: 217).[17] Wanano, an Eastern Tukanoan language of Brazil, assigns aspiration to original plain voiceless stops at beginning of initial syllables (and a few other environments; Waltz 2002).[18] Similarly, both Proto-Armenian and Proto-Germanic develop aspiration in the original Indo-European *T series. Many Armenian dialects have gone one step further, aspirating the Proto-Armenian plain voiceless series (descended from Indo-European plain voiced stops). Perhaps significantly, none of the 120 modern dialects surveyed in

should be [o] after [t] and [d], and therefore not susceptible to devoicing (which targets only high vowels), and (ii) if for some reason the normal rules of epenthesis were violated and [i] or [u] was inserted after a [t] or [d], this should trigger affrication (either [tʃi dʒi] or [tsu dzu]), which it does not.

[17] Note that innovation of aspiration in a non-aspirating area such as the Greek Islands is a better diagnostic for laryngeal markedness than is losing aspiration in a non-aspiration Sprachbund, such as happened with Dutch. The loss of Germanic aspiration in Dutch can easily be attributed to areal pressure from surrounding Romance languages, none of which aspirates voiceless stops.

[18] It is not clear whether this innovation is language-internal or the result of areal influence; Waltz (2002: 162) notes that Tariana, the Arawak language spoken by the tribe whose Periquitos group consider the Wanano their preferential marriage partners, has a contrastive series of voiceless aspirated stops.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

Jahukyan (1972) has deaspirated the Proto-Armenian voiceless aspirate series. Finally, in Southern Swahili the original voiceless plain and aspirated series merge as aspirated (Wald 1987: 997); the same happened in Standard Western Armenian and many other Western dialects (Vaux 1998b). These historical shifts from an original unaspirated series to an aspirated series (often as the result of a neutralisation process) are consistent with the hypothesis advanced in this paper, but difficult to explain if unaspirated stops are unmarked relative to their aspirated counterparts.

3.6 Interim conclusions concerning the unmarkedness of aspiration

In §3 we considered both internal and external evidence for the relative unmarkedness of voiceless aspirated stops *vis-à-vis* voiced stops and plain voiceless stops. The strongest internal evidence comes from phonological distributions, such as the elsewhere status of aspiration in English voiceless stops, and from neutralisation in first and second language phonological systems. External evidence for the relative unmarkedness of aspiration is ample; we have considered a number of relevant cases from acquisition, perception, production and language change.

4 Conclusions

Pulling back even further to the larger goals of this paper, we have suggested that the traditional belief that the maximally unmarked stop series in invariably voiceless and unaspirated (T) is not correct. Closer examination of phonological, phonetic, acquisitional, neural, aphasic and historical evidence suggests that:

(i) the maximally unmarked single-series stop is not voiceless and unaspirated, but rather is unspecified for laryngeal features, and hence can and does vary in its implementation;

(ii) the unmarked/more learnable two-way laryngeal opposition is (or at least can be) between unaspirated and aspirated stops (T:T^{h}), with the boundary between the two set at the English (long-lag) value.

A larger consequence of this proposal is that we cannot simply state that a given segment type is more or less marked than its counterpart based on its featural composition; we must instead consider the structure of the system of oppositions as a whole and perhaps positional considerations as well (cf. Battistella 1990: 132, Givón 1995: 27). This idea of course is not new, owing its core to Saussure, Jakobson and Trubetzkoy, but it has recently been insightfully elaborated by Rice (1999), who states (*contra* e.g. Lombardi 1995) that:

(15) a. Markedness cannot be defined in a single way.
b. Standard diagnostics for markedness, implication and frequency, are, by and large, unrevealing, leaving phonological patterning as the major evidence for markedness relations.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

c. Inventories play a role in determining markedness: depending upon the contrasts within a feature class, different features can emerge as unmarked with respect to submergence of the unmarked. (Cf. also Flemming 2005.)

d. Variation in markedness exists: two languages with identical surface contrasts within a class may have different unmarked features with respect to both submergence and emergence of the unmarked.

We have also made use at several points in this paper of a neo-Darwinian notion of markedness (Vaux 2002b), in which systems and segments that are more learnable are predicted to be more common cross-linguistically over evolutionary time. It is this dynamic, rather than some unspecified activity of an abstract notion of markedness, that accounts for the oft-invoked distributional data.

REFERENCES

Alber, Birgit (2001). Regional variation at edges: glottal stop epenthesis and dissimilation in Standard and Southern varieties of German. *Zeitschrift für Sprachwissenschaft* **20**. 3–41.

Alderete, John, Jill Beckman, Laura Benua, Amalia Gnanadesikan, John McCarthy & Suzanne Urbanczyk (1999). Reduplication with fixed segmentism. *LI* **30**. 327–364.

Allen, George D. & Sarah Hawkins (1980). Phonological rhythm: definition and development. In Yeni-Komshian *et al*. (1980: vol. 1). 227–256.

Altenberg, Evelyn P. & Robert M. Vago (1983). Theoretical implications of an error analysis of second language phonology production. *Language Learning* **33**. 427–447.

Aslin, Richard N., David B. Pisoni & Peter W. Jusczyk (1983). Auditory development and speech perception in infancy. In Marshall M. Haith & Joseph J. Campos (eds.) *Infancy and developmental psychobiology*. New York: Wiley. 573–687.

Auzou, Pascal, Canan Ozsancak, Richard J. Morris, Mary Jan, Francis Eustache & Didier Hannequin (2000). Voice onset time in aphasia, apraxia of speech and dysarthria: a review. *Clinical Linguistics and Phonetics* **14**. 131–150.

Avery, Peter (1997). Voicing specification and the specification of [voice]. *Toronto Working Papers in Linguistics* **16**. 21–35.

Avery, Peter & William J. Idsardi (forthcoming). *Laryngeal phonology*. Cambridge: Cambridge University Press.

Bach, Emmon (1991). Representations and operations in Haisla phonology. *WCCFL* **4**. 29–35.

Barna, Domonkos (1998). *Representation of consonants with unary features: aspiration in English*. MA thesis, Eötvös Loránd University, Budapest.

Battistella, Edwin L. (1990). *Markedness: the evaluative superstructure of language*. Albany: State University of New York Press.

Baum, Shari R. & Laurie Ryan (1993). Rate of speech effects in aphasia: voice onset time. *Brain and Language* **44**. 431–445.

Beach, Elizabeth, Denis Burnham & Christine Kitamura (2001). Bilingualism and the relationship between perception and production: Greek/English bilinguals and Thai bilabial stops. *International Journal of Bilingualism* **5**. 221–235.

Bernhardt, Barbara & Joseph Stemberger (1998). *Handbook of phonological development from the perspective of constraint-based nonlinear phonology*. San Diego: Academic Press.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

Bhatia, Tej K. (1976). On the predictive role of the recent theories of aspiration. *Phonetica* **33**. 62–74.
Blevins, Juliette (1993). Klamath laryngeal phonology. *IJAL* **59**. 237–279.
Blevins, Juliette (1995). The syllable in phonological theory. In John A. Goldsmith (ed.) *The handbook of phonological theory*. Cambridge, Mass. & Oxford: Blackwell. 206–244.
Broselow, Ellen, Su-I Chen & Chilin Wang (1998). The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition* **20**. 261–280.
Burling, Robbins (1959). Language development of a Garo and English speaking child. *Word* **15**. 45–68.
Burzio, Luigi (2000). Cycles, non-derived-environment blocking, and correspondence. In Joost Dekkers, Frank van der Leeuw & Jeroen van de Weijer (eds.) *Optimality Theory : phonology, syntax, and acquisition*. Oxford: Oxford University Press. 47–87.
Carr, J. (1953). An investigation of the spontaneous speech sounds of five-year-old deafborn children. *Journal of Speech and Hearing Disorders* **18**. 22–29.
Catts, Hugh & P. Jensen (1983). Speech timing of phonologically disordered children: voicing contrast of initial and final stop consonants. *Journal of Speech and Hearing Research* **26**. 501–510.
Cho, Young-mee Yu & Tracy Holloway King (2003). Semisyllables and universal syllabification. In Caroline Féry & Ruben van de Vijver (eds.) *The syllable in Optimality Theory*. Cambridge: Cambridge University Press. 183–212.
Chomsky, Noam & Morris Halle (1968). *The sound pattern of English*. New York: Harper & Row.
Clark, John & Colin Yallop (1995). *An introduction to phonetics and phonology*. 2nd edn. Oxford & Cambridge, Mass.: Blackwell.
Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook* **2**. 225–252.
Clements, G. N. (2005). The role of features in speech sound inventories. In Charles Cairns & Eric Raimy (eds.) *Phonological representations and architecture*. Cambridge, Mass.: MIT Press.
Coleman, John (2003). Discovering the acoustic correlates of phonological contrasts. *JPh* **31**. 351–372.
Compton, Arthur (1976). Generative studies of children's phonological disorders: clinical ramifications. In Donald M. Morehead & Ann E. Morehead (eds.) *Normal and deficient child language*. Baltimore: University Park Press. 61–96.
Cooper, Franklin S., Pierre C. Delattre, Alvin M. Liberman, John M. Borst & Louis J. Gerstman (1952). Some experiments on the perception of synthetic speech sounds. *JASA* **24**. 597–606.
Coyos, Jean-Baptiste (1994). Des occlusives aspirées en basque. *La Linguistique* **30:2**. 131–138.
Crowley, Terry (1983). Uradhi. In R. M. W. Dixon & Barry J. Blake (eds.) *Handbook of Australian languages*. Vol. 3. Amsterdam: Benjamins. 307–428.
Cruttenden, Alan (1985). *Language in infancy and childhood*. 2nd edn. Manchester: Manchester University Press.
Czaykowska-Higgins, Ewa & Marie Louise Willett (1997). Simple syllables in Nxaʔamxcín. *IJAL* **63**. 385–411.
Davenport, Mike & S. J. Hannahs (1998). *Introducing phonetics and phonology*. London & New York: Arnold.
Davis, Katharine (1995). Phonetic and phonological contrasts in the acquisition of voicing: voice onset time production in Hindi and English. *Journal of Child Language* **22**. 275–305.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

Deuchar, Margaret & Angeles Clark (1996). Early bilingual acquisition of the voicing contrast in English and Spanish. *JPh* **24**. 351–365.
Diehl, Randy L., A. F. Souther & C. L. Convis (1980). Conditions on rate normalization in speech perception. *Perception and Psychophysics* **27**. 435–443.
Dinnsen, Daniel A. (1992). Variation in developing and fully developed phonetic inventories. In Charles A. Ferguson, Lise Menn & Carol Stoel-Gammon (eds.) *Phonological development: models, research, implications*. Timonium, Md.: York Press. 191–210.
Dinnsen, Daniel A. (1999). Some empirical and theoretical issues in disordered child phonology. In William Ritchie & Tej K. Bhatia (eds.) *Handbook of child language acquisition*. New York: Academic Press. 647–674.
Disner, Sandra Ferrari (1983). *Vowel quality: the relation between universal and language-specific factors*. PhD dissertation, UCLA.
Dixit, Prakash (1987). Mechanisms for voicing and aspiration. *UCLA Working Papers in Phonetics* **67**. 49–102.
Dooling, Robert, Kazuo Okanoya & Susan Brown (1989). Speech perception by budgerigars (melopsittacus undulatus): the voiced–voiceless distinction. *Perception and Psychophysics* **46**. 65–71.
Dwyer, Arienne (2000). Consonantalization and obfuscation. In Aslı Göksel & Celia Kerslake (eds.) *Studies on Turkish and Turkic languages*. Wiesbaden: Harrassowitz. 423–439.
Dwyer, Arienne (2001). Predicting linguistic change: reweighing social, historical, and internal factors in a Sprachbund. Paper presented at the 2001 Annual Meeting of the German Linguistic Society, Leipzig.
Eckman, Fred R. (1981). On the naturalness of interlanguage phonological rules. *Language Learning* **31**. 195–216.
Edge, Beverly A. (1991). The production of word-final voiced obstruents in English by L1 speakers of Japanese and Cantonese. *Studies in Second Language Acquisition* **13**. 377–393.
Eguchi, S. & I. J. Hirsh (1969). Development of speech sounds in children. *Acta Oto-laryngologica Supplementum* **257**. 1–51.
Eilers, Rebecca (1980). Infant speech perception: history and mystery. In Yeni-Komshian *et al*. (1980: vol. 2). 23–39.
Eilers, Rebecca & Fred Minifie (1975). Fricative discrimination in early infancy. *Journal of Speech and Hearing Research* **18**. 158–167.
Eilers, Rebecca, D. Kimbrough Oller & Carmen R. Benito-Garcia (1984). The acquisition of voicing contrasts in Spanish and English learning infants and children: a longitudinal study. *Journal of Child Language* **11**. 313–336.
Eimas, Peter D., Einar R. Siqueland, Peter Jusczyk & James Vigorito (1971). Speech perception in infants. *Science* **171**. 303–306.
Elson, Ben (1947). Sierra Popoluca syllable structure. *IJAL* **13**. 13–17.
Engstrand, Olle & Karen Williams (1996). VOT in stop inventories and in young children's vocalizations: preliminary analyses. *Speech, Music, and Hearing Quarterly Progress and Status Report* **37:2**. 97–99.
Fellbaum, Marie (1996). Variation in the acquisition of voiceless stops in the interlanguage of second language learners of English and Spanish. *Proceedings of the 4th International Conference on Spoken Language Processing*. Available (November 2005) at http://www.asel.udel.edu/icslp/cdrom/vol3/663/a663.pdf.
Ferguson, Charles A. & Carol B. Farwell (1975). Words and sounds in early language acquisition. *Lg* **51**. 419–439.
Firchow, Irwin & Jacqueline Firchow (1969). An abbreviated phoneme inventory. *Anthropological Linguistics* **11**. 271–276.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

Fischer-Jørgensen, Eli (1980). Temporal relations in Danish tautosyllabic CV sequences with stop consonants. *Annual Report of the Institute of Phonetics, University of Copenhagen* **14**. 207–261.
Fischer-Jørgensen, Eli & Hajime Hirose (1974). A preliminary electromyographic study of labial and laryngeal muscles in Danish stop consonant production. *Haskins Laboratories Status Report on Speech Research* **SR-39/40**. 231–254.
Flege, James Emil, Martin J. McCutcheon & Steven C. Smith (1987). The development of skill in producing word final English stops. *JASA* **82**. 433–447.
Flemming, Edward (2005). Contrast and perceptual distinctiveness. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.) *Phonetically based phonology*. Cambridge: Cambridge University Press. 232–276.
Forrest, Karen & Barbara K. Rockman (1988). Acoustic and perceptual analysis of word-initial stop consonants in phonologically disordered children. *Journal of Speech and Hearing Research* **31**. 449–459.
Foster, Mary Lecron (1969). *The Tarascan language*. Berkeley & Los Angeles: University of California.
Foulkes, Paul, Gerry Docherty & Dominic Watt (1999). Tracking the emergence of structured variation realisations of (t) by Newcastle children. *Leeds Working Papers in Linguistics* **7**. 1–23.
Fujimura, O. (1971). Remarks on stop consonants: synthesis experiments and acoustic cues. In L. L. Hammerich, Roman Jakobson & Eberhard Zwirner (eds.) *Form and substance : phonetic and linguistic papers presented to Eli Fischer-Jørgensen*. Odense: Akademisk Forlag. 221–232.
Furbee-Losee, Louanna (1976). *The correct language – Tojolabal : a grammar with ethnographic notes*. New York & London: Garland.
Gandour, Jack, Soranee Holasuit Petty, Rochana Dardanaranda, Sumalee Dechongkit & Sunee Mukngoen (1986). The acquisition of the voicing contrast in Thai: a study of voice onset time in word-initial stop consonants. *Journal of Child Language* **13**. 561–572.
Gee, James Paul (1993). *An introduction to human language : fundamental concepts in linguistics*. Englewood Cliffs: Prentice Hall.
Gierut, Judith A. & Daniel A. Dinnsen (1986). On word-initial voicing: converging sources of evidence in phonologically disordered speech. *Language and Speech* **29**. 97–114.
Gilbert, John H. V. (1977). A voice onset analysis of apical stop production in 3-year-olds. *Journal of Child Language* **4**. 103–110.
Givón, T. (1995). *Functionalism and grammar*. Amsterdam & Philadelphia: Benjamins.
Gnanadesikan, Amalia (1997). *Phonology with ternary scales*. PhD dissertation, University of Massachusetts, Amherst.
Goad, Heather & Kathleen Brannen (2000). Syllabification at the right edge of words: parallels between child and adult grammars. *McGill Working Papers in Linguistics* **15**. 1–26.
Goad, Heather & Hyun-Sook Kang (2003). Word-final syllabification in L2 acquisition with emphasis on Korean learners of English. In Juana M. Liceras, Helmut Zobl & Helen Goodluck (eds.) *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference*. Somerville, Mass.: Cascadilla Press. 122–129.
Goldstein, Louis (2001). Online lecture notes from Ling 165b, 'Development of phonology'. Available (November 2005) at http://sapir.ling.yale.edu/ling165/order/order.html.
Golla, Victor (1970). *Hupa grammar*. PhD dissertation, University of California, Berkeley.
Gordon, Matthew (1998). The phonetics and phonology of non-modal vowels: a cross-linguistic perspective. *BLS* **24**. 93–105.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

Greenberg, Joseph H. (1966). *Language universals, with special reference to feature hierarchies*. The Hague & Paris: Mouton.
Grierson, George (1906). *The Piśâca languages of North-Western India*. London: The Royal Asiatic Society. Reprinted 1969, Delhi: Munshiram Manoharlal.
Haggard, Mark, Stephen Ambler & Mo Callow (1970). Pitch as a voicing cue. *JASA* **47**. 613–617.
Hale, Mark & Charles Reiss (1998). Formal and empirical arguments concerning phonological acquisition. *LI* **29**. 656–683.
Halle, Morris & Kenneth N. Stevens (1971). A note on laryngeal features. *MIT Quarterly Progress Report* **101**. 198–212.
Harris, John (1994). *English sound structure*. Oxford: Blackwell.
Harris, John (1996). Phonological output is redundancy-free and fully interpretable. In Jacques Durand & Bernard Laks (eds.) *Current trends in phonology : models and methods*. Salford: ESRI. 305–332.
Heyer, Sarah C. (1986). *English final consonants and the Chinese learner*. MA thesis, Southern Illinois University.
Holt, Lori L., Andrew J. Lotto & Randy L. Diehl (2004). Auditory discontinuities interact with categorization: implications for speech perception. *JASA* **116**. 1763–1773.
Holton, Gary (2000). *The phonology and morphology of the Tanacross Athabaskan language*. PhD dissertation, University of California, Santa Barbara.
Houlihan, Kathleen (1977). On aspiration and deaspiration processes. In Fred R. Eckman (ed.) *Current themes in linguistics : bilingualism, experimental linguistics, and language typologies*. Washington, D.C.: Hemisphere. 215–239.
Howe, Darin (2000). *Oowekyala segmental phonology*. PhD dissertation, UBC.
Hsu, Chai-Shune & Sun-Ah Jun (1998). Prosodic strengthening in Taiwanese: syntagmatic or paradigmatic? *UCLA Working Papers in Phonetics* **96**. 69–89.
Hume, Elizabeth (2003). Language specific markedness: the case of place of articulation. *Studies in Phonetics, Phonology and Morphology* **9**. 295–310.
Hume, Elizabeth & Georgios Tserdanelis (2002). Labial unmarkedness in Sri Lankan Portuguese Creole. *Phonology* **19**. 441–458.
Hutters, Birgit (1985). Vocal fold adjustments in aspirated and unaspirated stops in Danish. *Phonetica* **42**. 1–24.
Hyman, Larry (2001). On the limits of phonetic determinism in phonology: *NC revisited. In Elizabeth Hume & Keith Johnson (eds.) *The role of speech perception in phonology*. San Diego: Academic Press. 141–185.
Ioup, Georgette & Steven H. Weinberger (eds.) (1987). *Interlanguage phonology : the acquisition of a second language sound system*. Cambridge, Mass.: Newbury Press.
Irwin, Orvias C. (1947). Infant speech: consonantal sounds according to manner of articulation. *Journal of Speech Disorders* **12**. 402–404.
Iverson, Gregory K. & Sang-Cheol Ahn (2001). Post-obstruent tensing reconsidered. *Harvard Studies in Korean Linguistics* **9**. 107–116.
Iverson, Gregory K. & Joseph C. Salmons (1995). Aspiration and laryngeal representation in Germanic. *Phonology* **12**. 369–396.
Iverson, Gregory K. & Joseph C. Salmons (1999). Glottal spreading bias in Germanic. *Linguistische Berichte* **178**. 135–151.
Iverson, Gregory K. & Joseph C. Salmons (2003). Laryngeal enhancement in early Germanic. *Phonology* **20**. 43–74.
Jahukyan, Gevork^{h} (1972). *Haj barbařagitut^{h}jan Neratsut^{h}jun*. [*Introduction to Armenian dialectology*.] Erevan: Press of the Soviet Armenian Academy of Sciences.
Jakobson, Roman (1941). *Kindersprache, Aphasie, und allgemeine Lautgesetze*. The Hague: Mouton.
Jakobson, Roman & Morris Halle (1956). *Fundamentals of language*. The Hague: Mouton.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

Jäncke, Lutz (1994). Variability and duration of voice onset time and phonation in stuttering and nonstuttering adults. *Journal of Fluency Disorders* **19**. 21–37.
Jannedy, Stefanie (1995). Gestural phasing as an explanation for vowel devoicing in Turkish. *Ohio State University Working Papers in Linguistics* **45**. 56–84.
Jansen, Wouter (2004). *Laryngeal contrast and phonetic voicing : a laboratory phonology approach to English, Hungarian, and Dutch*. PhD dissertation, University of Groningen.
Jensen, John T. (1977). *Yapese reference grammar*. Honolulu: University Press of Hawaii.
Johansson, Victoria, Merle Home & Sven Strömqvist (2001). Final aspiration as a phrase boundary cue in Swedish: the case of att 'that'. *Working Papers, Phonetics Laboratory, Department of Linguistics, Lund University* **49**. 78–81.
Kari, James (1990). *Ahtna Athabaskan Dictionary*. Fairbanks: Alaska Native Language Center.
Keating, Patricia A. (1984). Phonetic and phonological representation of stop consonant voicing. *Lg* **60**. 286–319.
Keating, Patricia A., Wendy Linker & Marie Huffman (1983). Patterns in allophone distribution for voiced and voiceless stops. *JPh* **11**. 277–290.
Kenstowicz, Michael & Charles Kisseberth (1979). *Generative phonology : description and theory*. New York: Academic Press.
Kessinger, Rachel H. & Sheila E. Blumstein (1997). Effects of speaking rate on voice-onset time in Thai, French, and English. *JPh* **25**. 143–168.
Kewley-Port, Diane & Malcolm S. Preston (1974). Early apical stop production: a voice onset time analysis. *JPh* **2**. 195–210.
Khattab, Ghada (2000). VOT production in English and Arabic bilingual and monolingual children. *Leeds Working Papers in Linguistics* **8**. 95–122.
Kim, Chin-Wu (1970). A theory of aspiration. *Phonetica* **21**. 107–116.
Kiparsky, Paul (1979). Metrical structure assignment is cyclic. *LI* **10**. 421–441.
Kirchner, Robert (1994). Lenition in phonetically-based Optimality Theory. Ms, UCLA.
Kluender, Keith R., Randy L. Diehl & Peter R. Killeen (1987). Japanese quail can learn phonetic categories. *Science* **237**. 1195–1197.
Koenig, Laura L. (2001). Distributional characteristics of VOT in children's voiceless aspirated stops and interpretation of developmental trends. *Journal of Speech, Language, and Hearing Research* **44**. 1058–1068.
Kohler, Klaus J. (1984). Phonetic explanation in phonology: the feature fortis/lenis. *Phonetica* **41**. 150–174.
Kopkalli, Handan (1993). *A phonetic and phonological analysis of final devoicing in Turkish*. PhD dissertation, University of Michigan.
Kreidler, Charles W. (1989). *The pronunciation of English : a course book in phonology*. Oxford & New York: Blackwell.
Krueger, John R. (1962). *Yakut manual*. Bloomington: Indiana University.
Kuhl, Patricia K. & James D. Miller (1975). Speech perception by the chinchilla: voiced–voiceless distinction in alveolar plosive consonants. *Science* **190**. 69–72.
Kuhl, Patricia K. & James D. Miller (1978). Speech perception by the chinchilla: identification functions for synthetic VOT stimuli. *JASA* **63**. 905–917.
Kuhl, Patricia K. & Denise M. Padden (1982). Enhanced discriminability at the phonetic boundaries for the voicing feature in macaques. *Perception and Psychophysics* **32**. 542–550.
Labov, William, Malcah Yaeger & Richard Steiner (1972). *A quantitative study of sound change in progress*. Philadelphia: U.S. Regional Survey.
Lasky, Robert E., Ann Syrdal-Lasky & Robert F. Klein (1975). VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology* **20**. 215–225.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

Laver, John (1994). *Principles of phonetics*. Cambridge: Cambridge University Press.
Leopold, Werner F. (1939). *Speech development of a bilingual child: a linguist's record*. Vol. 1: *Vocabulary growth in the first two years*. Evanston & Chicago: Northwestern University.
Leopold, Werner F. (1947). *Speech development of a bilingual child: a linguist's record*. Vol. 2: *Sound learning in the first two years*. Evanston: Northwestern University Press.
Leslau, Wolf (1995). *Reference grammar of Amharic*. Wiesbaden: Harrassowitz.
Li, Bing (1996). *Tungusic vowel harmony: description and analysis*. The Hague: Holland Academic Graphics.
Liberman, A. M., P. C. Delattre & F. S. Cooper (1958). Some cues for the distinction between voiced and voiceless stops in initial position. *Language and Speech* **1**. 153–167.
Liljencrants, Johan & Björn Lindblom (1972). Numerical simulation of vowel quality systems: the role of perceptual contrast. *Lg* **48**. 839–862.
Lindblom, Björn (1986). Phonetic universals in vowel systems. In John J. Ohala & Jeri J. Jaeger (eds.) *Experimental phonology*. Orlando: Academic Press. 13–44.
Lisker, Leigh (1975). Is it VOT or a first-formant transition detector? *JASA* **57**. 1547–1551.
Lisker, Leigh (1977). Factors in the maintenance and cessation of voicing. *Phonetica* **34**. 304–306.
Lisker, Leigh (1984). How is the aspiration of English /p, t, k/ 'predictable'? *Language and Speech* **27**. 391–394.
Lisker, Leigh & Arthur S. Abramson (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word* **20**. 384–422.
Local, John (2002). Variable domains and variable relevance: interpreting phonetic exponents. Paper presented at the Aix–Cambridge Symposium on Speech Perception, Aix-en-Provence.
Lombardi, Linda (1991). *Laryngeal features and laryngeal neutralization*. PhD dissertation, University of Massachusetts, Amherst. Published 1994, New York: Garland.
Lombardi, Linda (1995). Laryngeal features and privativity. *The Linguistic Review* **12**. 35–59.
McCarthy, John J. (1988). Feature geometry and dependency: a review. *Phonetica* **45**. 84–108.
McCarthy, John J. (2000). Harmonic serialism and parallelism. *NELS* **30**. 501–524.
McCarthy, John J. & Alan Prince (1994). The emergence of the unmarked: optimality in prosodic morphology. *NELS* **24**. 333–379.
Macken, Marlys A. (1980). Aspects of the acquisition of stop systems: a crosslinguistic perspective. In Yeni-Komshian *et al*. (1980: vol. 1). 143–168.
Macken, Marlys A. & David Barton (1979). The acquisition of the voicing contrast in Spanish: a phonetic and phonological study of word-initial stop consonants. *Papers and Reports on Child Language Development* **16**. 42–66.
Macken, Marlys A. & David Barton (1980). The acquisition of the voicing contrast in English: a study of voice onset time in word-initial stop consonants. *Journal of Child Language* **7**. 41–74.
MacNeilage, Peter F. (1980). The control of speech production. In Yeni-Komshian *et al*. (1980: vol. 1). 9–21.
Maddieson, Ian (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.
Maddieson, Ian, Caroline L. Smith & Nicola Bessell (2001). Aspects of the phonetics of Tlingit. *Anthropological Linguistics* **43**. 135–176.
Major, Roy C. & Eunyi Kim (1999). The similarity differential rate hypothesis. *Language Learning* **49**. Suppl. 1. 151–183.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

Marlow, Patrick E. (2000). Koyukon sound system and orthography. In Jules Jetté & Eliza Jones. *Koyukon Athabaskan dictionary*. Fairbanks: Alaska Native Language Center. lxvi–lxxi.
Masica, Colin (1991). *The Indo-Aryan languages*. Cambridge: Cambridge University Press.
Menn, Lise (1976). *Pattern, control, and contrast in beginning speech : a case study in the development of word form and word function*. PhD dissertation, University of Illinois, Champaign-Urbana. Distributed 1978, Indiana University Linguistics Club.
Menyuk, Paula & Mary Klatt (1975). Voice onset time in consonant cluster production by children and adults. *Journal of Child Language* **2**. 223–231.
Mikola, Tibor (1996). Zur Geschichte der auslautenden Tenues im Lappischen. In Lars-Gunnar Larsson (ed.) *Lapponica et Uralica : 100 Jahre finnisch-ugrischer Unterricht an der Universität Uppsala*. Uppsala: University of Uppsala. 99–102.
Miller, Joanne L. (1981). Some effects of speaking rate on phonetic perception. *Phonetica* **38**. 159–180.
Miller, Joanne L. & Lydia E. Volaitis (1989). Some effects of speaking rate on the perceived internal structure of phonetic categories. *Perception and Psychophysics* **46**. 505–512.
Morris, Richard J. & W. S. Brown, Jr. (1994). Age-related differences in speech variability among women. *Journal of Communication Disorders* **27**. 49–64.
Morton, Katherine & Mark Tatham (1980). Devoicing, aspiration, and nasality: cases of universal misunderstanding? *Occasional Papers, Department of Language and Linguistics, University of Essex* **23**. 90–103.
Mulford, Randa & Barbara Hecht (1980). Learning to speak without an accent: acquisition of a second-language phonology. *Papers and Reports on Child Language Development* **18**. 16–74.
Nathan, Geoffrey S., Warren Anderson & Dudsaba Budsayamongkon (1987). On the acquisition of aspiration. In Ioup & Weinberger (1987). 204–212.
Nevin, Bruce E. (1998). *Aspects of Pit River phonology*. PhD dissertation, University of Pennsylvania.
Newton, Brian (1972). *The generative interpretation of dialect : a study of modern Greek phonology*. Cambridge: Cambridge University Press.
Odden, David (1987). Dissimilation as deletion in Chukchi. *Proceedings of the Eastern States Conference on Linguistics* (*ESCOL*) **4**. 235–246.
Odden, David (1996). Review of Spencer (1996). *Phonology* **13**. 446–449.
Odden, David (2001). Introduction to phonological analysis. Ms, Ohio State University.
Olmsted, D. L. (1971). *Out of the mouth of babes : earliest stages in language acquisition*. The Hague & Paris: Mouton.
Orton, Harold & Eugen Dieth (1962–1971). *Survey of English dialects*. Leeds: Arnold.
Özsancak, Canan, Pascal Auzou, M. Jan, S. Léonardon, M. Gaillard & D. Hannequin (1997). Pseudobulbar and Parkinsonian dysarthria: an acoustical analysis. *Journal of Neurology* **244**. 53.
Pater, Joe (1999). Austronesian nasal substitution and other NÇ effects. In René Kager, Harry van der Hulst & Wim Zonneveld (eds.) *The prosody–morphology interface*. Cambridge: Cambridge University Press. 310–343.
Phillips, Colin (2001). Levels of representation in the electrophysiology of speech perception. *Cognitive Science* **25**. 711–731.
Pierrehumbert, Janet (2000). The phonetic grounding of phonology. *Les Cahiers de l'ICP : Bulletin de la Communication Parlée* **5**. 7–23.
Pisowicz, Andrzej (1976). *Le développement du consonantisme arménien*. Wrocław: Ossolineum.
Pisowicz, Andrzej (1987). Developmental tendencies in the voiced–voiceless opposition of stop consonants (on the basis of data from Persian and Armenian).
https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

In Stanisław Piłaszewicz & Jerzy Tulisow (eds.) *Problems of Asian and African languages*. Warsaw: Polish Academy of Sciences. 233–238.
Poedjosoedarmo, Gloria (1996). Variation and change in the sound systems of Brunei dialects of Malay. In Peter W. Martin, Conrad Ožóg & Gloria Poedjosoedarmo (eds.) *Language use and language change in Brunei Darussalam*. Athens, Oh.: Ohio University Press. 37–42.
Polomé, Edgar C. (1967). *Swahili language handbook*. Washington: Center for Applied Linguistics.
Poon, Pamela G. & Catherine A. Mateer (1985). A study of VOT in Nepali stop consonants. *Phonetica* **42**. 39–47.
Port, Robert F. & Michael L. O'Dell (1985). Neutralization of syllable-final voicing in German. *JPh* **13**. 455–471.
Prator, Clifford & Betty Wallace Robinett (1972). *Manual of American English pronunciation*. 3rd edn. New York: Holt, Rinehart & Winston.
Prentice, D. J. (1971). *The Murut languages of Sabah*. Canberra: Australian National University.
Preston, Martin, Grace Yeni-Komshian & Rachel Stark (1967). Voicing in initial stop consonants produced by children in the pre-linguistic period from different language communities. *Annual Report, Neurocommunications Laboratory, Johns Hopkins University School of Medicine* **1**. 307–323.
Prince, Alan (1997). Endogenous constraints on Optimality Theory. Paper presented at the Hopkins Optimality Theory Workshop/University of Maryland Mayfest 1997, Baltimore.
Prince, Alan & Paul Smolensky (1993). *Optimality Theory: constraint interaction in generative grammar*. Ms, Rutgers University & University of Colorado, Boulder. Published 2004, Malden, Mass. & Oxford: Blackwell.
Pyun, Kyung-ah (1999). Syllable based influences on the interlanguage phonology of Korean ESL learners. Ms, Arizona State University.
Rédei, Károly (1965). *Northern Ostyak chrestomathy*. The Hague: Mouton.
Regnier, Susan (1993). Quiegolani Zapotec phonology. *Work Papers of the Summer Institute of Linguistics, University of North Dakota* **37**. 37–63.
Reiss, Charles (1997a). Review of Barbara Bernhardt, John Gilbert & David Ingram (eds.) *Proceedings of the UBC International Conference on Phonological Acquisition*. Somerville, Mass.: Cascadilla Press. *Linguist List 8.1545*.
Reiss, Charles (1997b). Re: question of functionalism. *Linguist List 8.690*.
Repp, Bruno H. (1979). Relative amplitude of aspiration noise as a voicing cue for syllable-initial stop consonants. *Language and Speech* **22**. 173–189.
Rice, Keren (1999). Featural markedness in phonology: variation. *Glot International* **4:7**. 3–6. **4:8**. 3–7.
Ringen, Catherine & Petur Helgason (2002). Voice and spread glottis in Swedish. Paper presented at the 76th Annual Meeting of the Linguistic Society of America, San Francisco.
Rose, Yvan (2000). *Headedness and prosodic licensing in the L1 acquisition of phonology*. PhD dissertation, McGill University.
Ross, Steven (1994). The ins and outs of paragoge and apocope in Japanese–English interphonology. *Second Language Research* **10**. 1–24.
Schadeberg, Thilo C. (1999). Katupha's Law in Makhuwa. In Jean-Marie Hombert & Larry M. Hyman (eds.) *Bantu historical linguistics: theoretical and empirical perspectives*. Stanford: CSLI. 379–394.
Schiefer, Lieselotte (1992). Trading relations in the perception of stops and their implications for a phonological theory. In Gerard J. Docherty & D. Robert Ladd (eds.) *Papers in laboratory phonology II: gesture, segment, prosody*. Cambridge: Cambridge University Press. 296–313.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

Scobbie, James M. (2002). Flexibility in the face of incompatible English VOT systems. Paper presented at the 8th Conference on Laboratory Phonology (Labphon 8), Yale University.
Scobbie, James M., Fiona Gibbon, William J. Hardcastle & Paul Fletcher (2000). Covert contrast as a stage in the acquisition of phonetics and phonology. In Michael B. Broe & Janet B. Pierrehumbert (eds.) *Papers in laboratory phonology V : acquisition and the lexicon*. Cambridge: Cambridge University Press. 194–207.
Silverman, Daniel (1995). Optimal and sub-optimal timing relations between stops, vowels, and vocal fold spreading. Abstract of paper presented at the 26th Meeting of the North East Linguistic Society, Harvard University.
Silverman, Daniel (1998). Alveolar stops in American English and the nature of allophony. *NELS* **28**. 425–435.
Singhaniyom, Jitravadee (1999). *The relationship between the perception of Thai stops and the degree of voice-timing in the production of English stops by native speakers of English*. MA thesis, Chulalongkorn University.
Slowiaczek, Louisa M. & Daniel A. Dinnsen (1985). On the neutralizing status of Polish word-final devoicing. *JPh* **13**. 325–341.
Smith, Bruce L. (1978). Temporal aspects of English speech production: a developmental perspective. *JPh* **6**. 37–67.
Smith, Caroline L. (1997). The devoicing of /z/ in American English: effects of local and prosodic context. *JPh* **25**. 471–500.
Smith, Neil (1973). *The acquisition of phonology : case study*. Cambridge: Cambridge University Press.
Smolensky, Paul (1996). On the comprehension/production dilemma in child language. *LI* **27**. 720–731.
So, L. K. H. & B. J. Dodd (1994). Phonologically disordered Cantonese-speaking children. *Clinical Linguistics and Phonetics* **8**. 235–255.
Sohn, Hyang-Sook (2001). Optimization of word-final coronals in Korean loanword adaptation. In Caroline Féry, Antony Dubach Green & Ruben van de Vijver (eds.) *Proceedings of HILP 5*. Potsdam: University of Potsdam. 159–177.
Sorianello, Patrizia (1997). Per una ridefinizione del processo di aspirazione nei dialetti calabresi. *Studi Italiani di Linguistica Teorica e Applicata* **26**. 81–103.
Spencer, Andrew (1996). *Phonology : theory and description*. Oxford & Cambridge, Mass.: Blackwell.
Spencer, Robert F. (1946). The phonemes of Keresan. *IJAL* **12**. 229–236.
Stampe, David (1978). *A dissertation on Natural Phonology*. New York & London: Garland.
Steele, Jeffrey (2002). L2 learners' modification of target language syllable structure: prosodic licensing effects in interlanguage phonology. In A. James & J. Leather (eds.) *New Sounds 2000 : Proceedings of the 4th International Symposium on the Acquisition of Second-language Speech*. Klagenfurt: University of Klagenfurt. 315–324.
Steriade, Donca (1997). Phonetics in phonology: the case of laryngeal neutralization. Ms, UCLA.
Steriade, Donca (2001). The phonology of perceptibility effects: the P-map and its consequences for constraint organization. Ms, UCLA.
Stevens, Kenneth N. (1972). The quantal nature of speech: evidence from articulatory-acoustic data. In Edward E. David & Peter B. Denes (eds.) *Human communication : a unified view*. New York: McGraw-Hill. 51–66.
Stevens, Kenneth N. (1989). On the quantal nature of speech. *JPh* **17**. 3–45.
Stevens, Kenneth N. & Sheila E. Blumstein (1981). The search for invariant acoustic correlates of phonetic features. In Peter D. Eimas & Joanne L. Miller (eds.) *Perspectives on the study of speech*. Hillsdale: Erlbaum. 1–38.
Stevens, Kenneth N. & Dennis H. Klatt (1974). Role for formant transitions in the voiced–voiceless distinction for stops. *JASA* **55**. 643–659.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

Straight, H. Stephen (1976). *The acquisition of Maya phonology : variation in Yucatec child language*. New York & London: Garland.
Streeter, Lynn A. (1976). Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature* **259**. 39–41.
Summerfield, A. Quentin & Mark P. Haggard (1974). Perceptual processing of multiple cues and contexts: effects of following vowel upon stop consonant voicing. *JPh* **2**. 279–295.
Summerfield, A. Quentin & Mark P. Haggard (1977). On the dissociation of spectral and temporal cues to the voicing distinction in initial stop consonants. *JASA* **62**. 435–448.
Sweeting, P. M. & R. J. Baken (1982). Voice onset time in a normal-aged population. *Journal of Speech and Hearing Research* **25**. 129–134.
Syeed, Syed Mohammad (1978). The Himalayan way of breathing the last: on the neutralizing status of the word-final aspiration in Kashmiri. *The Eastern Anthropologist* **31**. 531–541.
Szigetvári, Péter (1996). Laryngeal contrasts and problematic representations. *The Even Yearbook* **2**. 97–110.
Taff, Alice, Lorna Rozelle, Taehong Cho, Peter Ladefoged, Moses Dirks & Jacob Wegelin (2001). Phonetic structures of Aleut. *JPh* **29**. 231–271.
Tatham, Mark (1979). Some problems in phonetic theory. In Harry Hollien & Patricia Hollien (eds.) *Current issues in the phonetic sciences*. Amsterdam: Benjamins. 93–106.
Tatham, Mark & Katherine Morton (1971). An EMG experiment refuting the high-level segmental context-sensitive theory. Ms, University of Essex.
Trubetzkoy, N. S. (1958). *Grundzüge der Phonologie*. 2nd edn. Göttingen: van der Hoeck & Ruprecht.
Tsui, Ida Y. H. (1996). Perception of aspiration and place of articulation of initial stops by normal and sensorineural hearing-impaired listeners. Ms, University of Hong Kong.
Tsui, Ida Y. H. & Valter Ciocca (2000). Perception of aspiration and place of articulation of Cantonese initial stops by normal and sensorineural hearing-impaired listeners. *International Journal of Language and Communication Disorders* **35**. 507–525.
Tyler, Ann A. & John Saxman (1991). Initial voicing contrast acquisition in normal and phonologically disordered children. *Applied Psycholinguistics* **12**. 453–479.
Tyler, Ann A. & K. H. Watterson (1991). Effects of phonological versus language intervention in preschoolers with both phonological and language impairment. *Child Language Teaching and Therapy* **7**. 141–160.
Vaux, Bert (1998a). English aspiration, glottalization, and flapping revisited. Ms, Harvard University.
Vaux, Bert (1998b). *The phonology of Armenian*. Oxford: Clarendon.
Vaux, Bert (2001). Consonant insertion and hypercorrection. Paper presented at the 75th Annual Meeting of the Linguistic Society of America, Washington, D.C.
Vaux, Bert (2002a). Aspiration in English. Paper presented at the University of Wisconsin-Madison.
Vaux, Bert (2002b). The evolution of vowel systems: Dispersion Theory *vs*. evolution. Paper presented at the North American Phonology Conference, Concordia University, Montreal.
Vihman, Marilyn May (1996). *Phonological development : the origins of language in the child*. Cambridge, Mass. & Oxford: Blackwell.
Vijayakrishnan, K. G. (1999). Weakening processes in the Optimality framework. Available as ROA-329 from the Rutgers Optimality Archive.
Viswanath, Nagalapura & David Rosenfield (2000). Preponderance of lead voice onset times in stutterers under varying constraints. *Communication Disorders Quarterly* **22**. 49–55.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press

Wagner, Michael (2002). The role of prosody in laryngeal neutralization. *MIT Working Papers in Linguistics* **42**. 357–376.
Wald, Benji (1987). Swahili and the Bantu languages. In Bernard Comrie (ed.) *The world's major languages*. London & Sydney: Croom Helm. 991–1014.
Waltz, Nathan E. (2002). Innovations in Wanano (Eastern Tucanoan) when compared to Piratapuyo. *IJAL* **68**. 157–215.
Waters, R. & W. Wilson (1976). Speech perception by rhesus monkeys: the voicing distinction in synthesized labial and velar stop consonants. *Perception and Psychophysics* **19**. 285–289.
Weinberger, Steven H. (1987). The influence of linguistic context on syllable simplification. In Ioup & Weinberger (1987). 401–417.
Wells, John C. (1982). *The accents of English*. 3 vols. Cambridge: Cambridge University Press.
Westbury, John R. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *JASA* **73**. 1322–1336.
Westbury, John R. & Patricia A. Keating (1986). On the naturalness of stop consonant voicing. *JL* **22**. 145–166.
Whalen, D. H., Catherine T. Best & Julia R. Irwin (1997). Lexical effects in the perception and production of American English /p/ allophones. *JPh* **25**. 501–528.
Williams, Lee (1977). The voicing contrast in Spanish. *JPh* **5**. 169–184.
Williams, Lee (1980). Phonetic variation as a function of second-language learning. In Yeni-Komshian *et al*. (1980: vol. 2). 185–215.
Wilson, Colin (2001). Consonant cluster neutralisation and targeted constraints. *Phonology* **18**. 147–197.
Winitz, Harris, Conrad LaRiviere & Eve Herriman (1975). Variations in VOT for English initial stops. *JPh* **3**. 41–52.
Yavaş, Mehmet (1994). Final stop devoicing in interlanguage. In Mehmet Yavaş (ed.) *First and second language phonology*. San Diego: Singular Publishing Group. 267–282.
Yavaş, Mehmet (1997). The effects of vowel height and place of articulation in interlanguage final stop devoicing. *International Review of Applied Linguistics* **35**. 115–125.
Yeni-Komshian, Grace H., James F. Kavanagh & Charles A. Ferguson (eds.) (1980). *Child phonology*. 2 vols. New York: Academic Press.
Yu, Alan (2004). Explaining final obstruent voicing in Lezgian: phonetics and history. *Lg* **80**. 73–97.

https://doi.org/10.1017/S0952675705000667 Published online by Cambridge University Press