Acoust. Sci. & Tech. **26**, 5 (2005)

INVITED REVIEW

The marriage of phonetics and phonology

John J. Ohala

Department of Linguistics, University of California, Berkeley 94720, USA

Keywords: Phonetics, Phonology, Nasalization, Stops, Fricatives

PACS number: 43.70.–h, 43.70.Bk [DOI: 10.1250/ast.26.418]

1. INTRODUCTION

The distinction usually made between phonetics and phonology is that phonetics studies the physical or physiological aspects of speech, including its articulatory, aerodynamic, acoustic, auditory, and perceptual aspects, whereas phonology is concerned with accounting for the variation in speech sounds in different but related languages and dialects, and within a given language in the environment of different morphemes, different positions within an utterance, word, or other speech sounds. Phonology also derives generalizations about particular languages and common patterns of speech sounds in all human languages. More modern manifestations of phonology purport to characterize the psychological or even the genetic underpinnings of language, including its sounds. Table 1 presents some examples of generalizations that might be deemed exclusively the domain of phonetics (in the first column) or phonology (in the second column).

In the earliest known descriptions of speech sounds in languages there was no distinction between phonetics and phonology. Panini, for example, who worked in approximately the 5th c. BPE, compiled a magnificent and detailed description of Sanskrit speech sounds and their contextual variations [1]. The beginnings of the differentiation between phonetics and phonology, however, probably began in the 19th century and accelerated in the 20th. A number of cumulative developments account for this:

- The 19th century saw the success of historical linguistics in establishing family relationships between languages by discovering systematic phonetic relationships between numerous words in two or more languages. Table 2 gives some examples. This permitted Rask [2], for example, to extract the generalizations equivalent to 'voiced stops in Latin and Greek corresponds to a voiceless stop articulated at the same place in Germanic languages' (such as English).
- In 1861–62 Schleicher [3] posited abstract parent forms—marked with an asterisk—for cognate words in the parent language, now called Proto-Indo-European (see column 4 in Table 2).
- In 1878 Ferdinand de Saussure [4] posited the existence of sounds (now known as 'laryngeals') in Proto-Indo-European, these sounds being known only by their effects on other sounds since they had disappeared in the languages on which Saussure based his study (Subsequently some of the posited 'laryngeals' were found in Hittite texts).
- In the early 20th c. Saussure [5] taught that beside the immediately observable facts of a language's sounds there was also an underlying reality, its structure or system, that needed to be established.
- The concept of the phoneme arose from the teachings and practice of Baudouin de Courtenay [6], Saussure [5], and Sweet [7], the phoneme, which might have a variety of contextually-determined phonetic variants, was regarded as the minimally contrastive unit in the make-up of words and morphemes. To give one simple example: the initial sound in "pip" [pʰɪp] and the second sound in "spill" [spɪl] were said to be members of the same phoneme /p/ (identified as a phoneme by the use of the forward slashes); the first member being aspirated and the second one not. Conceptually the phoneme is parallel to the one in historical linguistics where, e.g., a single parent sound *p is reconstructed for the variants found in related languages, i.e., [p] in Latin and [f] in Germanic (see above). The phoneme was conceived of as the psychological 'parent' form of the contextually-determined phonetic variants, e.g., a phoneme /h/ in English was the psychological source of the phonetic variants [h] in "how" [hɑu] and [ç] in "hue" [çju].
- With the rise of generative phonology [8] variant forms of morphemes such as *profane* /pʰrofejn/ ~ *profanity* /pʰrofænɪri/ were presumed to be derived from a common abstract underlying form /profǣnɪti/ in the mental lexicon. Thus it was claimed that a

Table 1 Examples of phonetic and phonological descriptions.

Phonetic description	Phonological description
[tʰ], the initial sound in the English word *time*, is a voiceless, aspirated, alveolar stop.	The initial obstruents in German *fuss* [fus] "foot" and Latin *ped* "foot" derive from a *p, a voiceless labial stop, in the parent language, Proto-Indo-European
French [ɛ] has a higher 2nd formant (F_2) than French [a].	The initial fricative in French *chien* /ʃjɛ̃/ "dog" derives from an earlier /k/, still found in the Picardy dialect where this word is pounced /kjɛ̃/.
The burst of dorsal stops like [k] contain a compact spectral peak that is lacking in stops made at further forward places of articulation	The stressed vowels in English *extreme* [ɪkˈstrim] and *extremity* [ɪkˈstrɛmɪɾi] are derived (via the psychological grammar possessed by the native speaker) from an underlying [eː].
The vowel [ɚ], as in American English *lurk* [lɚk] has the lowest F_3 of all vowels, achieved by three simultaneous constrictions in the vocal tract: at the lips, in the mid-palatal region, and in the pharynx.	English does not permit a labial or velar consonant (those characterized by the Jakobsonian feature [+grave]) to appear after the diphthong /ɑʊ/; only alveolar or palatal consonants are permitted in this position: *loud* [lɑʊd], *out* [ɑʊt], *couch* [kʰɑʊt͡ʃ]
The high rate of listener's confusion of [p] and [kʷ] is due to their having similar formant transitions, differing perhaps only in the rate of transition.	In many languages there is constraint forbidding intervocalic heterorganic clusters, thus Late Latin **noktu** *night* became Italian **notte**.

Table 2 Cognate words in various Indo-European languages demonstrating that they all originated from a common parent language which has disappeared (Transcription simplified).

Greek	Latin	English	Proto-Indo-European
pod	ped	foot	*ped
phrater	frater	brother	*bhrater-
pherei	fero	bear (verb)	*bher-
duo	duo	two	*dwo
(gi-)gnoskein	(ko-)gnoskere	know	*gno-

careful study of sound patterns could reveal psychological facts.

Bit by bit, linguists got accustomed to the idea that there could be ever more abstract representations and processes for speech sounds which could account for their behavior and which were distinct from phonetic representations and processes.

Today phonology is often practiced as a virtually autonomous enterprise, divorced from phonetics and other empirically-oriented disciplines. Most phonologists feel free to posit increasingly abstract structures whose relationship to the real world—including the psychological domain—may be questioned [9–11]. There is, however, a movement to improve the empirical base of phonology [12,13].

Whatever may happen in phonology in the domain of theory, it has always been at least an inductive discipline in that it focuses on general patterns of speech sound variation, diachronically and synchronically. As such it still produces a virtual treasure of information about the behavior of speech sounds and this treasure can be exploited by researchers in speech technology and speech pathology. In the remaining sections of this paper I will focus on the kinds of characteristic variations in pronunciation that phonology has documented and suggest how they can be explained by reference to physical phonetic factors. I work with the following assumption: sound change can arise when the listener misinterprets the speech signal—either because of ambiguities created by the speaker or because of the listener's own inattention or failure to hear speech sounds that are weak or obscured by noise.

2. EMERGENT STOPS

Many speech sounds are produced where a specific cavity in the vocal tract has at least two potential exit valves, i.e., conduits to the space outside the speaker for the transmission of sound and/or the venting of the airflow that is essential to speech. This is represented schematically in Fig. 1. Buccal sounds (those made at or further forward than the uvular place of articulation) have the oral and the velic valve as two potential exit channels. Laterals like [l, ʎ] exploit two exit valves: that manifested at the tongue midline, which is closed, and that manifested by at least one channel at the side of the tongue (often including the buccal sulcus, the space between the teeth and the cheek), which valve is open. When speech sounds which are sequenced one after the other which have opposite valvular configurations, that is, valve A closed and valve B open in one sound (as shown in the top of Fig. 1) and the reverse in the other sound (shown at the middle of Fig. 1), there is some probability that in the transition between the two sounds both valves may be closed thus creating a stop (as shown in the bottom of Fig. 1). These are commonly called 'epenthetic' stops but I prefer the term 'emergent' stops,

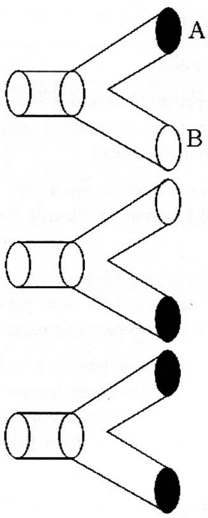

Fig. 1 Schematic representation of how emergent stops may arise in the transition between two sounds neither of which are themselves stops.

Table 3 Emergent stops in sequences of nasal+fricative/liquid.

Engl.	youngster [jʌŋjˈ**k**stɚ]	<	jʌŋ + stɚ
Engl.	warmth [wɔɻm**p**θ]	<	*warm* + θ
Engl.	Thom**p**son	<	*Thom* + *son*
Engl.	dem**p**ster 'judge'	<	*deem* + *ster*
Sotho	voni**t**ʃa 'to show'	<	voniʃa (caus. 'to see')
Cl. Gk	an**d**ros	<	*anĕros* 'man'
French	cham**b**re	<	Lat. *kamĕra* 'room'
Span	alham**b**ra	<	Ar. *al hamra* 'the red'
Latin	tem**p**lum	<	**tem - lo* 'a section'

Table 4 Emergent stops between laterals and adjacent apical fricatives.

Engl.	else [ɛl**t**s]			
	Ilse [ɪl**t**sə]			
Kwakiutl	k′wēɬ**t**so	<	k′wēɬ + so	'to be feasted'
Greek	hes**t**hlos	<	heslos	
Ital.	Ischia [is**k**ja] <		is**k**la < is**t**la < isla 'island'	
	schiavo [s**k**jiavo] < *s**k**lavo < *s**t**lavo < slavo 'slave'			
	(the ultimate source of *ciao!* [tʃao])			

using the word 'emergent' as in evolutionary biology to describe some novel entity that 'emerges' from a rearrangement or adaptation of pre-existing elements. Table 3 provides some examples of stops emerging in the transition between a nasal and a following oral segment and Table 4 of a stop emerging between a lateral and medial fricative.

In the case of *Ischia* (the name of an island off the coast of Naples, Italy) and *schiavo*, the emergent [t] was replaced by a [k] and then the [l] was replaced by a palatal glide via sound changes that are independent of the change that yielded the emergent [t].

A somewhat unusual instance of stop emergence (conjectured by Ohala [14]) involves the overlap of two buccal closures which may be part of a sequence where neither one is an obstruent. In this case the stop that is formed must be a type of click. A click is made by trapping air between two buccal closures, increasing the volume of this cavity by lowering the tongue. In this case the two exit valves to this chamber are the labial and the alveolar valves. Table 5 gives some examples cited by Ohala [14].

Table 5 Emergent labial click ([ʘ]) between a sequence of [m] followed by [n].

Middle Engl.	nem**p**ne	< nemʘne OEngl. < namna 'name'	
	som**p**nour	< somʘnour < somnour 'summoner'	
Landais Fr.	dam**p**nadge	< damʘnaticu < damnaticu 'damning'	
	fem**p**ne	< femʘna < femna < femina 'woman'	
Old Swed.	häm**p**na	< hämʘna < hämna 'revenge'	

3. EMERGENT FRICATIVES AND AFFRICATION

In English and Japanese the voiceless glottal fricative /h/ is phonetically just a voiceless copy of the following sound which in both languages must be a sonorant, i.e., a non-obstruent sound. The allophone of /h/ that appears before the voiced palatal sonorants is the voiceless palatal fricative [ç], e.g., English /hju/ "hue" is phonetically [çju]; Japanese /hito/ is [çito], /hjɑku/ is [çjɑku]. (On the other hand, the allophone of /h/ before other vowels is simply a voiceless glottal fricative, e.g., English "hat" /hæt/ is phonetically [hæt] which is equivalent to [æ̥æt].) How can a voiceless glottal fricative become a voiceless palatal fricative? The answer is that frication is generated when air flows rapidly through a narrow orifice. During a voiceless /h/, which has a wide-open glottis, there is a high rate of airflow and the narrowest point through which the air flows is in the lingual-palatal region where the palatal glide /j/ or the palatal vowel /i/ has a very narrow constriction. The other vowels do not have such a narrow constriction so that even if their configuration is anticipated during the /h/ no appreciable supra-glottal noise will be generated.

Emergent frication can also be found when stops, especially voiceless apical stops, are released before high close vowels or glides like [i] or [j], as exemplified in (1) and (2). Table 6 gives data on the development of affricated stops where the earlier forms (Proto-Bantu) shows that the process crucially depends on the following vowel being a high one like [i] and [u]; the affrication does not occur if the following vowels are the less high vowels, like [ɪ o ɑ]

Table 6 The development of affricated stops from stops followed by high close vowels, but not when the following vowel is less high (The '*' marks forms reconstructed by comparing cognate words in dozens of "sister" Bantu languages).

Proto-Bantu	***Mvumbo***	***Translation***
*-bumɑ	b͡vumo	*fruit*
*-dib-	d͡ʒiwo	*shut*
*-kubɑ	p͡fuwo	*chicken*
*-tiitʊ	t͡ʃir	*animal*
BUT:		
*-bod	buo	*spoil (v.)*
*-dɪ	di	*eat*
*-gɑdɑ	kɑlɑ	*mat*
*-konde	kwɑnde	*banana*

(1) English *actual* [æktʃuəlˠ] < aekt + juəl
bestial [bistʃəlˠ] < bist + jəl
(2) Latin fortia > *fortsja > fortse > Fr. forse ([15], p. 130)

4. NASALIZATION AFFECTING F_1

In many languages the nasalization of a vowel has the effect of changing the quality of the vowel, principally the height of the vowel. This is seen, for example, in the French cognates *final* [finɑl] ~ *fin* [fɛ̃], *brune* [bʁyn] ~ *brun* [bʁœ̃]. Electrical analogues of the vocal tract reveal that among the acoustic effects of nasalizing a vowel there is an increase in F_1 (and concomitant increase in bandwidth) [16]. Thus this sound pattern exemplified in French can be partially explained since there is an inverse relationship between perceived vowel height and F_1; see Fig. 2 (based on [16]).

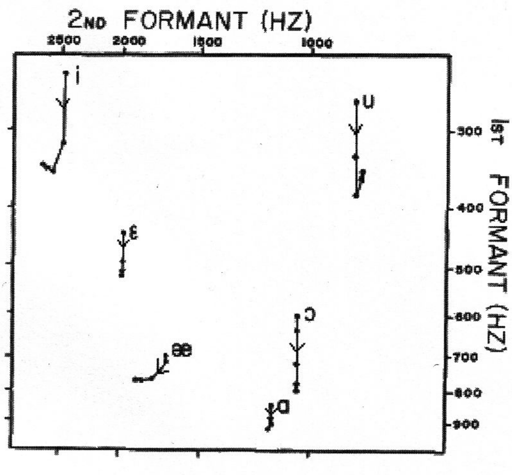

Fig. 2 The changes in F1 and F2 for selected vowels as a function of increasing coupling with the nasal cavity (the direction of the arrows show incremental increases in nasal coupling); based on data from House & Stevens (1956).

5. VELAR STOPS BECOME APICAL AFFRICATES BEFORE PALATALS

One of the most common sound changes encountered in many different languages is the change from a velar stop to an apical stop, most often with simultaneous affrication of the apical stop when this stop appears before a palatal vowel like [i] or the palatal approximant [j]. Examples are English cool /kul/ ~ chill /t͡ʃɪl/ and the above mentioned dialectal variation in French, /kjɛ̃/ ~ /ʃjɛ̃/ "dog" (see Table 1). This can be accounted for by reference to phonetic factors. The burst and formant transitions of the phonetic sequences [ki] or [kj] on the one hand and [t] on the other, are acoustically very similar. In fact, in a perceptual study of isolated CV syllables where C = [p], [t], or [k] and V = [i], [ɑ], or [u], the highest rate of confusion, 38%, occurred with [ki] which was misidentified as [ti] [17]. The fricative release in the above-cited sound changes can be explained by aerodynamic factors: the long, narrow channel created by the tongue against the palate and the necessarily high rate of airflow through this channel when the stop is released are ideal circumstances for the generation of audible turbulence (see above).

Moreover, this sound change and the confusion patterns in the Winitz *et al.* study is asymmetrical: whereas /ki/ is confused as /ti/ 38%, the confusion of /ti/ as /ki/ occurred only 3% of the time (in the most relevant condition). This can also be explained by reference to acoustic-perceptual facts: the releases in /ki/ and /ti/ are overall highly similar but differ largely in the fact that that /ki/ has a narrow bandwidth peak in the noise spectrum (around 3 kHz)—this is essentially the front cavity resonance—whereas the /ti/ lacks this feature. Stop bursts are important cues to stop place. It is plausible to assume that failure to detect such a distinctive acoustic feature is more likely than imagining its presence when it is absent. Therefore if this mid-frequency peak is not detected then the stop burst will sound like that of /t/. The reverse confusion, /ti/ taken as /ki/ would require the unlikely perceptual mistake of "adding" this missing spectral peak to the burst's spectrum (see [18,19]).

6. CONCLUSION

Phonetics can provide some of the explanations for the sound patterns discovered by phonologists [20]. But the benefit is symmetrical: phonology, by studying the behavior of speech sounds in language—sound changes, patterns of sound sequences in words and morphemes, allophonic variation, the structure of phoneme inventories—can help phonetics to focus on those articulatory, acoustic, and perceptual factors that principally serve the function of communication.

REFERENCES

[1] W. S. Allen, *Phonetics in Ancient India* (Oxford University Press, London, 1953).

[2] R. Rask, *Undersøgelse om det gamle Nordiske eller Islandske Sprogs Oprindelse* (Gyldendalske Boghandlings Forlag, Copenhagen, 1818).

[3] A. Schleicher, *Compendium der vergleichenden Grammatik der ndo-germanischen Sprachen* (Weimar, 1861–1862).

[4] F. de Saussure, *Mémoire sur le système primitif des voyelles dans les langues indo-européennes* (Vieweg, Paris, 1878).

[5] F. de Saussure, *Cours de linguistique générale* (Payot, Paris, 1916).

[6] J. Baudouin de Courtenay, *Versuch einer Theorie phonetischer Alternationen. Ein Capitel aus der Psychophonetik* (Karl J. Trübner, Strassburg, 1895).

[7] H. Sweet, *A Primer of Phonetics* (Clarendon Press, Oxford, 1890).

[8] N. Chomsky and M. Halle, *The sound Pattern of English* (Harper & Row, New York, 1968).

[9] J. J. Ohala, "There is no interface between phonetics and phonology. A personal view," *J. Phonet.*, **18**, 153–171 (1990).

[10] J. J. Ohala, "The integration of phonetics and phonology," *Proc. XIIth Int. Congr. Phonetic Sciences*, Aix-en-Provence, 19–24 Aug. 1991, Vol. 1, pp. 1–16 (1991).

[11] J. J. Ohala, "The relation between phonetics and phonology," in *Handbook of Phonetics*, W. Hardcastle and J. Laver, Eds. (Blackwell's, Oxford, 1996), pp. 674–694.

[12] J. Kingston and M. E. Beckman, Eds., *Papers in Laboratory Phonology I: Between Grammar and Physics of Speech* (Cambridge University Press, Cambridge, 1990).

[13] J. J. Ohala and J. J. Jaeger, Eds. *Experimental Phonology* (Academic Press, Orlando, 1986).

[14] J. J. Ohala, "A probable case of clicks influencing the sound patterns of some European languages," *Phonetica*, **52**, 160–170 (1995).

[15] M. K. Pope, *From Latin to Modern French with Especial Consideration of Anglo-Norman* (Manchester University Press, Manchester, 1934).

[16] A. House and K. Stevens, "Analog studies of the nasalization of vowels," *J. Speech Hear. Res.*, **21**, 218–232 (1956).

[17] H. Winitz, M. E. Scheib and J. A. Reeds, "Identification of stops and vowels for the burst portion of /p, t, k/ isolated from conversational speech," *J. Acoust. Soc. Am.*, **51**, 1309–1317 (1972).

[18] J. J. Ohala, "Comparison of speech sounds: Distance vs. cost metrics," in *Speech Production and Language*, S. Kiritani, H. Hirose and H. Fujisaki, Eds. (Mouton de Gruyter, Berlin, 1997), pp. 261–270.

[19] M. C. Plauché, C. Delogu and J. J. Ohala, "Asymmetries in consonant confusion," *Proc. Eurospeech*, Rhodes, Greece, 22–25 Sep. 1997, pp. 2187–2190 (1997).

[20] J. J. Ohala, "The phonetics of sound change," in *Historical Linguistics: Problems and Perspectives*, C. Jones, Ed. (Longman, London, 1993), pp. 237–278.