

A critical introduction to phonology: Of sound, mind, and body

Daniel Silverman

Part 3 I speak, therefore you are

Chapter 7 *Loquor ergo es*

In a paper published in 1911, Etienne Lombard reported on a series of experiments in which he found that subjects increase the volume of their voice when the ambient noise level increases.

This result has come to be known in some circles as the *Lombard Effect*. There have been a number of ways that the Lombard Effect has been interpreted. One straightforward interpretation is that the effect is an inherently *social* phenomenon: speakers maintain a speech-to-noise ratio that is favorable to communication, because if they don't, listeners might not hear them accurately. This can be interpreted as a form of *altruism* on the part of speakers, in that they want to ease the burden of their interlocutors. However, it can just as readily be interpreted as *selfish* in motivation: speakers want to effectively get their ideas across so that their own needs might be satisfied. Either way however, this *social* interpretation of the Lombard Effect assumes that speakers are making use of a *public* feedback loop, requiring feedback from *listeners* in order to adjust their own speaking volume.

But Lombard also found that when speakers' ability to monitor their *own* speech is diminished (by, say, wearing earplugs), also, they raise their voices. This suggests that the sound of one's own voice (called the *sidetone*) might be necessary in order to properly regulate its volume, and so there is not necessarily an inherent social dimension to the voice adjustment. Instead, the feedback loop may be wholly internal, or *private*, which is consistent with a *solipsistic* interpretation of the Lombard Effect. That is, speakers are not solely responding to feedback from the outside world, but instead, their behavior is a consequence of self-monitoring. These three interpretations are tabulated in Figure 7.1.

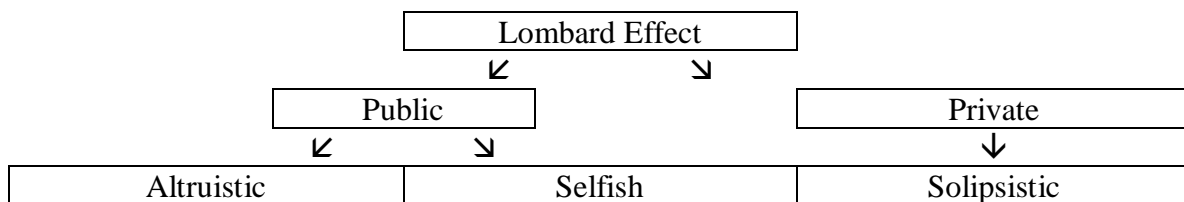


Figure 7.1. Three interpretations of the Lombard Effect.

Altruism, selfishness, and, more generally, speakers' *intentions*, have sometimes been invoked to account not only for the Lombard Effect, but also for phonological patterns in general. In this chapter I investigate only one example each of "altruistic" and

“selfish” approaches to phonology that have been explored in the literature, that are indeed characteristic—or, rather, symptomatic—of broad schools of linguistic thought. Unlike the Lombard Effect though, these effects aren’t presumed to hold only under abnormal conditions. Instead, they are assumed to be fully active under normal conditions, in the sense that our mastery of language is claimed to be due to active mental processes of a broadly “altruistic” and/or “selfish” nature; an assumption I reject. I will suggest that any approach to phonological structure which makes reference to speakers’ supposed *intentions* should be regarded with suspicion.

Most of these approaches, indeed most current phonological models in general, have their origins in Sapir’s highly influential paper of 1933 on the supposed “psychological reality of the phoneme”. (Recall from Chapter One that John Whitney’s “feeling a ‘t’” supposedly reflected an abstract sound for which there was no phonetic evidence; the phoneme.) For adherents of Sapir’s approach (fleshed out most fully by the linguists Noam Chomsky and Morris Halle in their book from 1968), alternants *derive* from a *single* underlying element, or phoneme. The underlying (or “input”) value *turns into* another value (or “output value”) as a consequence of a system of phonological rules or constraints. To take an example, adherents to this school of thought assume—arbitrarily, by most standards—that in cases of spirantization it is the stop that is underlyingly present (by convention, surrounded by virgules, as in Figure 7.2), and that this stop becomes a fricative in the relevant context, but stays a stop in others. For adherents to the Sapir/Chomsky and Halle approach then, the various alternants derive from a single psychological entity—the phoneme—which takes on new phonetic characteristics depending on its context.

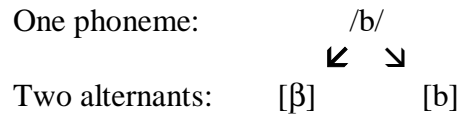


Figure 7.2. Two alternants derived from a single value.

This approach is very different from the one presented in this book. The difference hinges on my contention that *functional non-distinctness is not the same as identity of mental origin*. In the present approach, I have been capturing the functional non-distinctness of allophonic alternants with set notation: components of morphemes that alternate with each other emerge as functionally non-distinct. So recall an example from Corsican.

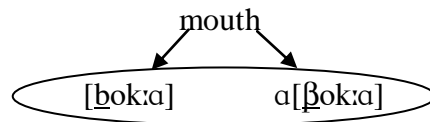


Figure 7.3. Two alternants whose functional identity is established due to their alternation.

In Corsican, there are (at least) two forms of the same word—[bok:a] and (a)[βok:a]. As learners come to realize that these two forms have the same meaning, the

phonetic differences between them are set into high relief against the stable phonetic background, and the functional non-distinctness between the two forms is established. Each allomorph—and so each *allophone*—is as psychologically “legitimate” or “authentic” as the other, and so neither has a more privileged mental status. Learners *do not* conclude that the [b] and [β] are actually derived from the same underlying building block, that is, /b/, simply because there is no evidence available that would prompt them to arrive at such a conclusion.

Revisiting elements of our discussion in Chapter Two, in this chapter I suggest that the phonemic approach to phonology can trace its origins to so-called “alphabetism”: because most linguists have been trained in an alphabetic orthography, their intuitions tell them that phonology is organized in accordance with alphabetic/phonemic principles. As will be shown, the uncontroversial mismatch between “phonemism” and the actual phonetic content of the speech signal has required the “phonemists” to posit a multi-leveled phonology. One level, the “underlying level” is phonemic in organization, while an additional level captures the superficial aspects of speech production. The phonemists focus their energies on uncovering the psychological route from the underlying string of phonemes to the superficial content of the speech signal.

In this chapter I suggest that, because of its multi-leveled characterization of phonology, the phonemic approach often demands a conflation of synchrony and diachrony, and also a conflation of phonetics and cognition. It further requires positing a dichotomy between speakers’ *intentions* and speakers’ *performance*. I reject this position in its entirety.

I conclude by reiterating that phonology is best characterized as a self-organized system of substantive social conventions which evolves passively over generations of speakers. The regularities we observe in phonological systems are due to a complex interaction of phonetic and cognitive pressures acting over generations and generations of language use, and can be understood only when considering the communicative function of language itself. Hence, *loquor ergo es* “I speak, therefore you are”.

EASE OF PERCEPTION

At least since the nineteenth century, linguists have entertained the proposal that phonological patterns may be influenced by a “struggle” between two opposing pressures: *ease of perception* and *ease of production*. Indeed, variants of this view have persisted in some theoretical circles up to the present day, and into the pages of this book. Regarding ease of perception, I have been providing evidence all along that the acoustic dispersion of contrastive sounds, and the acoustic quality of the transition from one articulatory posture to the next, are slowly and passively influenced by phonetic and functional pressures on speech variants that render communication *successful*. In a related way, allophonic alternations may be seen as the passive, long-term result of these same sorts of selectional pressures. The result is that linguistic signals are quite easy to perceive, but this is not due to the *intentions*, conscious or otherwise, of individual speakers.

ARE SPEAKERS “ALTRUISTIC”?

Although few would deny that the sounds of phonological systems are inherently easy to perceive by their possessors, the *origin* or *locus* of this inherent property is not

fully agreed upon. While I have been providing evidence for a *diachronic* origin and an inherently *social* locus of this property of language, at least some researchers have placed the locus elsewhere; square in the lap of individual speakers. Consider an example case. In a paper from 2002, the linguist John Kingston discussed a well-known property of contrastively nasalized low vowels (for example, [ã]): these vowels tend to have *more* nasalization (in the form of a more significantly lowered soft palate) than do mid or high nasalized vowels. The greater degree of nasalization here is probably related to the fact that *most* low vowels have some degree of soft palate lowering, even those that are not contrastively nasalized. This soft palate lowering seems to be an automatic articulatory concomitant of implementing a low vowel. Consequently, if a language is to possess a *contrast* in nasality on its low vowels (such as [a]-[ã]), then the nasalized form needs to be *especially* nasalized, in the form of *increased* soft palate lowering, so that it is rendered distinct from other sounds ([ã]-[ã̃]).

Readers of Chapters Five and Six know that all the components—speech variation, exemplar theory, probability matching—are in place to understand the diachronic origins of this pattern. Kingston however, places the locus of this mechanism elsewhere: “[S]peakers exert themselves to convey contrasts in ways that are entirely unexpected if they couldn’t optimize their pronunciations to ensure that contrasts are maintained”. He concludes that “Speakers must be altruists”.

This approach raises several questions. First, should we assume that speakers are being specifically *altruistic*? As noted in our discussion of the Lombard Effect, any “exertion” on the part of speakers to render speech in a perspicuous fashion may just as readily be interpreted as *selfish* in origin, and not altruistic. Speakers may find that *not* exerting themselves while speaking does not get them what they want. Consequently, they may exert themselves to speak clearly in order to satisfy their own needs. Of course, this interpretation of the facts of speech is just as suspect as an altruistic one, but it does point to the arbitrary nature of Kingston’s proposal. Perhaps, under the deepest genetic reading of the term, speakers may be regarded as altruistic in some sense, since linguistic evolution has *passively* resulted in a system of communication that is beneficial to listeners, much as genuine natural selection may have imbued the proper genetic material to produce certain altruistic behaviors in animals. However, in the synchronic sense suggested by Kingston—who, recall, suggests that speakers “exert themselves”—speakers are far from altruistic. They are simply reproducing what they perceive.

But Kingston’s synchronic, speaker-based locus of his proposed mechanism begs more fundamental questions. First, if one does not consider the *evolution* of such patterns, the extra soft palate lowering on low nasalized vowels is indeed “entirely unexpected” unless speakers “ensure” their speech is easier for listeners to interpret clearly. But it’s not clear why a pattern that has *diachronically* evolved to a particular state should be accounted for in *synchronic* terms. When diachrony is taken into account, the pattern isn’t “entirely unexpected” at all. In fact, it makes perfect, natural sense.

Second, once speakers are—by any reasonable description of the facts—*faithfully reproducing* the speech patterns they encounter, why would they need to apply an overlay of altruistic intent to “optimize their pronunciations so that contrasts are maintained”? Each speaker is *not* determining anew the articulatory configuration that might assist a listener (a listener who may have just been assisting this very speaker, by the way!) in recovering the contrastive sounds of the language. Rather, all speakers who *implement*

their low nasalized vowels with a greater degree of soft palate lowering have always *heard* these nasalized vowels with a greater degree of soft palate lowering. So again, speakers are simply reproducing what they hear.

Third, if we assume that certain aspects of phonological systems have their origins in an altruistic intent on the part of speakers, we must further assume a non-altruistic component of phonology as well, that is, a component of speech that is not a result of speakers' "exerting themselves" to speak more clearly. But since speakers *copy* the very speech patterns that they hear, and they do so remarkably accurately, where would the evidence lie—both for the speaker and for the linguist—that speech patterns are indeed divisible along altruistic and non-altruistic lines?

As I remarked in Chapter Three, and have elaborated upon in subsequent discussion, present-day phonological patterns result from a dizzyingly complex and long-term interaction of articulatory, acoustic, auditory, aerodynamic, perceptual, functional, and social forces, and it's facile to assume that the explanation for present-day phonetic and phonological patterns reduces to present-day factors. By failing to consider the evolution-like selectional pressures that give rise to speech patterns, and instead placing a synchronic altruistic burden on individual speakers, Kingston is playing an unwinnable guessing game about speakers' mental states. Hermann Paul, considering this sort of proposal in 1880, wrote, "...[T]here is no such thing as a conscious effort made to prevent a sound change. For those who are affected by the change have no suspicion that there is anything to guard against, and they habitually pass their lives in the belief that they speak today as they spoke years ago, and that they will continue to the end to speak in the same way." If we replace the word "conscious" with "unconscious" (which is no doubt the psychological level at which Kingston intends his "altruism" to be operating), Paul's remark serves as a befitting rejoinder to Kingston's proposals of over a century later.

In sum, phonological systems are indeed organized such that their sounds are easy to perceive by their possessors, but speakers, to put it starkly, are no more altruistic than are foraging rats and ducks.

EASE OF PRODUCTION

In addition to phonological systems being organized such that their sounds are easy to perceive, many linguists have also argued that phonological systems are organized such that their sounds are *easy to produce*. In her 2001 book, the linguist Joan Bybee makes use of this notion when she observes a link between the *frequency of usage* and the *tendency to simplify speech patterns*. As in the present approach, Bybee argues that linguistic categories emerge as a consequence of patterns' use and re-use. According to Bybee, if sound changes are the result of phonetic processes that apply as a consequence of actual language use, then those words that are used more frequently are more likely to undergo phonetic simplifications. She provides many case studies—mostly from English and Spanish—illustrating how sound changes that simplify speech patterns may begin with words and phrases of the highest frequency, and then may gradually come to influence the pronunciation of other, less frequently employed items. For example, frequent words like "camera" and "every" have lost their second vowels ([k^hæm.ɪ], [ɛv.ɪ]), whereas less common words with comparable structure retain these vowels: "mammary," "homily" ([ˈmæm.ə.ɹi], [ˈhɑm.ə.li]).

Bybee's research has its origins in the proposals of the nineteenth century scholar Hugo Schuchardt. In his monograph of 1885, Schuchardt considered a number of examples of articulatory simplification in frequent words such as titles and greetings. For example, Hungarian *alazatos szolgája* (a greeting like "your servant") has become *alá szolgáj*, Spanish *vuestra merced* has become *usted* ("you"), and German *guten Morgen* is often *g'Morgen* ("good morning"). He offers the following account of such simplifications:

The change of a sound, its progress in a certain direction... consists of the sum of microscopic displacements. It is, therefore dependent upon the number of repetitions. If *x* requires 10,000 repetitions to become *x'*, these repetitions are to be counted within individual words, nevertheless. An *x* spoken one time each in 10,000 different words would not become *x'*. I will not deny that a word that has been spoken 10,000 times can favor the development of the sound *x* to *x'* in a word spoken only 8000 times, etc. The greater or lesser frequency in the use of individual words... is... of great importance for their phonetic transformation... Rarely-used words drag behind; very frequently used ones hurry ahead... They have been compared to small coins that, as they pass from hand to hand rapidly, are soon worn thin.

In other words, frequent words may lead the way in a simplifying sound change, and due to their lead, increasingly less frequent words may be recruited in this simplification as well. Indeed, in our discussion of Trique in Chapter Five, we considered how passive pressures toward homophone avoidance may trigger individual words to undergo pioneering changes, and may, in turn, induce other words to follow.

I should note, however, that "ease of production" is not the same as "simplicity of production". With repetitive use, the words "mountain" and "Trenton" may *simplify* from [ˈmæ̃ntʰɪn] and [ˈtɹɛ̃ntʰɪn] to [ˈmæ̃ʔn] and [ˈtɹɛ̃ʔn]. But these latter pronunciations might not be *easier* to produce; just ask a non-native speaker of English to pronounce both variants, and see what happens!

At a presentation in 2002, the linguist Mark Hale criticized, or more properly, *parodied* the Schuchardt/Bybee approach to the relationship between frequency and simplification with the expression "practice makes imperfect": while piano players, for example, get *better and better* at their chosen task as they repeat and repeat the proper movements, language users, according to the parody, actually get *worse and worse*, since, according to Hale's characterization of the Schuchardt/Bybee approach, they become less and less adept at implementing all the necessary speech sounds. Hence, practice makes imperfect.

However, Hale does not consider the possibility that *exactly because* certain words are frequently encountered in the speech stream, they are more *predictably* present. For example, English has sequences like "have to" and "going to", which, because the meanings which underlie these constructions are needed quite often when expressing one's thoughts, are used over and over again. Because of their constant repetition and their consequent predictability, those particular spontaneous variants that are slightly simplified may yet effectively convey the intended meaning to listeners. Due to

probability matching, in time, these simplifications may become conventionalized, subject to further simplification, and “have to” ([ˈhæ:v̩ tʰu]) and “going to” ([ˈɡoʊɪŋ tʰu]) eventually become “hafta” ([ˈhæftə]), and “gonna” ([ˈɡʌnə]). Other constructions of increasingly less similar structure may consequently follow suit, depending on their frequency. So, by the year 1981, Joe Ely was able to release an album called “Musta Notta Gotta Lotta” (as in “I musta notta gotta lotta sleep last night”) and all speakers of American English can understand exactly what he means. Consequently, even with less phonetic information, the intended message may be effectively communicated to the listener: frequent simplified productions may be correctly interpreted by listeners, and might come to be used more often by speakers as a consequence of probability matching. By contrast, less frequent items are *less* predictable, and so require *more* phonetic information in order to be unambiguously communicated to listeners. These items will be more resistant to simplification, because those variants that are indeed simplified may not be successfully communicated, and so only non-simplified variants are more likely to be pooled with those tokens over which probabilities are matched.

Hale’s parody is based on the supposition that individual speakers are somehow constantly getting it *wrong*; that the locus of phonological simplification resides in speakers’ incorrectly implementing what they are *trying* to say, and doing so over and over again. But in fact, the typically *slow-acting* nature of phonological change—due in part to speakers’ demonstrated talent for probability matching—shows that speakers are constantly getting it remarkably *right*, and that the locus of Schuchardt’s and Bybee’s observed simplifications lies not in *speakers’ supposedly unreliable production of speech*, but instead in *listeners’ demonstrably reliable perception of meaning*.

Most fundamentally, Hale’s “practice makes perfect” parody suggests a marked disconnect between language function and language structure. Piano players practice and practice in order to faithfully and accurately copy what is written in black and white in a music score. In piano playing, copying the immutable score is the be-all and end-all of the process (apart from the interpretative component that separates the artist from the hack). The “practice makes imperfect” parody assumes that there is indeed an immutable black and white “linguistic score”—with a string phonemes taking the role of a string of musical notes—that speakers are striving to reproduce. But, as demonstrated in this book, language users are not attempting to read an immutable “linguistic score” the way a pianist does a musical one. Rather, speakers are intending to communicate *meaning* to listeners, and the speech signal is simply the *medium* by which this information is transmitted from speaker to listener. So, while it’s true that speakers are remarkably adept at faithfully copying the perceived speech signal—variation and all—this copying is clearly mediated by the accuracy with which listeners recover the *meaning* which underlies the acoustic content of the speech signal. Under those circumstances where meaning is recoverable *despite* variation in speech (as in Schuchardt’s and Bybee’s simplifications) or *because of* variation in speech (as discussed in Chapters Five and Six), then certain variants—even simplified variants—effectively serve their linguistic function, and may consequently become conventionalized.

Ultimately, Hale’s parody suggests a solipsistic view of language. Just like those researchers who might believe that the Lombard Effect is a consequence of a wholly internal feedback loop by which speakers modify their voice level solely as a consequence of self-monitoring, Hale too seems to assume that the social milieu plays no

significant role in how language comes to be cognitively represented, and that language structure is largely independent of language use, language experience, and language function. In truth, *language use, language experience, and language function influence language structure, and any attempt to deny this fact will not advance our understanding of the organizing principles of language.*

ARE SPEAKERS “SELFISH”?

A number of phonologists have recently taken a very bold position on the matter of ease-of-production in phonology. They propose that ease-of-production is not (or not *only*) a slow-going, *passive physical constraint* on phonology, but, rather, is an *active mental constraint* that is part of an abstract system of phonological knowledge that every speaker automatically brings to the task of language learning. The constraints may be of different strengths in different languages, but nonetheless, they are assumed to be provided to us automatically as a simple consequence of our being Homo sapiens. For these phonologists, the task of the language learner is to determine the strength of one constraint relative to others (the constraints themselves, recall, come “for free”, although their strengths vary on a language-specific basis); the task of the linguist is to discover the constraints, and like the learner, to determine their relative strengths for any given language. The particular family of abstract constraints relating to “effort minimization” has been termed “lazy” by one of the most enthusiastic proponents of this approach, the linguist Robert Kirchner. “Lazy” constraints are in conflict with “faithfulness” constraints, which place a premium on keeping speech as similar as possible to hypothesized *non-lazy* or *pre-lazy* mental representations of speech sounds; this is the level of the “alphabetic score”; the phonemic level. In a paper from 2004, Kirchner writes that every possible phonetic realization of every phonological structure is matched against a mental calculation of its “...effort cost, the biomechanical energy required for [its] articulatory production...” Linguistic forms which abide by the abstract constraint then, is one which does not exceed the calculated effort cost allotted by the mental calculation, while remaining as faithful to the underlying value as the faithfulness constraints demand. It is these forms—out of a potentially infinite set of candidate forms—that actually earn the privilege of coming out of our mouths.

As with Kingston’s “altruism”, if “laziness” is part and parcel of the linguistic system, listeners must somehow be able to partition the incoming speech signal into at least two distinct components: that component which is a consequence of speakers’ laziness, and the remainder, the component of the speech signal that is “faithful” to the “pre-lazy” phonemic level. These phonemic linguistic representations are neither *spoken* by speakers, nor *heard* by listeners, but language users, because of the constraint system that they inherit, come to figure out they do indeed exist.

Let’s briefly consider an example of how this works. As discussed at length in Chapter Six, many languages have an alternation between voiced stops and voiced fricatives, such that the fricative is found between vowels. For Kirchner, the stop and fricative alternants are phonologically the same entity, in that the stop is realized as a fricative due to the active laziness constraints; unlike the analysis presented in Chapter Six, Kirchner assumes that it is indeed more natural to produce an intervocalic fricative as opposed to an intervocalic stop, and so spirantization is a “lazy” realization of the stop. So at the most abstract, phonemic, or “underlying” level, the fricative is mentally

represented as a stop, but, due to the laziness constraints, it is realized as a fricative when it finds itself between vowels.

Kirchner is actually making two conflations in his approach to the phonological role of ease of production. He is conflating the phonetic with the cognitive, and, like Kingston, he is conflating the diachronic with the synchronic. Let's consider each of these conflations in turn.

First, Kirchner's boldest assertion is essentially that *purely physical, phonetic* pressures on linguistic sound systems—a supposedly quantifiable measure of “articulatory effort”—should be expressed in *purely abstract, psychological* terms. He writes in 2004 (using the common term *lenition* to refer to spirantization and certain other processes), “...I proceed from the intuition that lenition is driven by an imperative to minimize articulatory effort. Unlike standard approaches, however, I argue that lenition patterns arise directly from this effort minimization constraint (which I style LAZY)...”. For Kirchner then, demonstrable phonetic tendencies are fully “psychologized”, one might say; they are incorporated into a wholly abstract, psychological system of knowledge that every speaker possesses. The system of phonetically-rooted abstract constraints, then, is a synchronically active component of our linguistic knowledge.

This conflation of the phonetic and the cognitive is in stark contrast to the proposals espoused herein. Indeed, as Mikołaj Kruszewski wrote in 1881 (and as I have taken as the jumping-off point for this book),

Language occupies a completely isolated place in the realm of nature: it is a combination of physiological and acoustic phenomena governed by physical laws, and of unconscious and psychical phenomena governed by laws of an entirely different kind. This fact leads us to a most important question: what is the relation...between the physical principle and the unconscious and psychical principle?

Kirchner's answer to this question is that *phonetic* tendencies that might slowly play themselves out diachronically are recapitulated in (and conflated with) the synchronic *psychological* system that speakers are claimed to possess. However, it is difficult to see the advantage—or even the motivation—for this conflation of the *physical*—shaped by one set of laws and principles—and the *cognitive*—shaped by a wholly unrelated set of laws and principles. My misgivings here stem less from concerns about *parsimony* as such—indeed, most natural systems suffer from a harrowing degree of complexity—but rather, from concerns about *plausibility*: proponents of this general approach unfortunately eschew the obvious (and obviously unanswerable) question of *how and why* creeping, slow-going phonetic tendencies have, over the course of human evolution, been transformed into all-at-once mental constraints on linguistic sound structure. Evolutionarily speaking, it would seem difficult to isolate the environmental and genetic factors that might have given rise to this remarkable saltation.

Regarding the conflation of the *diachronic* and the *synchronic*, again, Kirchner's approach contrasts with the present one, which argues that sounds in alternation in the present have evolved in the absence of the users who come to possess them, that present-day alternations have no present-day causes; they only have present-day effects. Once again, I bow to Hermann Paul, who wrote in 1880:

One of the commonest errors is the supposition that a change which has arisen in a long period by numerous small displacements is to be referred to a single act resulting from a desire for convenience...The truth is that...the motory sensation developed by tradition...belong[s] to a period perhaps long and gone. It is equally mistaken to refer the appearance of a sound change in each case to some particular manifestation of laziness, weariness, or neglect, and to ascribe its non-appearance in other cases to some special care and observation.

In short, while there may be ample textual, comparative, and/or internal reconstructive evidence to conclude that certain fricative alternants *historically* derive from stops, there is no evidence to conclude that such fricatives *psychologically* derive from stops.

“ALPHABETISM” AND “PHONEMISM”: WHENCE AND WITHER

It is fascinating to read her scrupulous analysis of stage technique, where she notes that the name Fang is funny because it ends abruptly with a hard consonant...

-Jane and Michael Stern, reviewing Phyllis Diller's memoir, *My Life in Comedy*. New York Times Book Review, March 13, 2005.

How far has alphabetism led innocent language users down the garden path? Consider the quote above, and contemplate its multi-faceted mistakenness. New York Times reviewers Jane and Michael Stern find it fascinating that American comedian Phyllis Diller apparently thought the name “Fang” (her loosely fictional no-good rotten husband) was funny “because it ends abruptly with a hard consonant.” The “hard consonant” in question is presumably “g”, since that is the last letter in “Fang”. But you and I know that this is merely an orthographic convention; the English digraph “ng” actually represents the tongue-body nasal, [ŋ]: [fæŋ]. This nasal is actually one of the “softest” consonants you’re likely to encounter, since, if you recall from Chapter Two, it is very vowel-like in its acoustic characteristics. So Diller is wrong on that count. The Sterns got it wrong as well, since they find Diller’s “scrupulous analysis” of the word “Fang” so “fascinating”. If “Fang” really is funny, it certainly has nothing to do with the quality of its final consonant, at least according to the criteria that Diller establishes and that the Sterns buy into.

The wonderful comedian Buddy Hackett seemed to know better. He loved his name, because it had so many sounds that you could “hook on to”, that is, it had so many stops. Neil Simon got it right too, when, in *The Sunshine Boys*, the ex-Vaudevillian Willie Clark intones, “I’m in this business 57 years. I know what’s funny. Words with a ‘k’ in it are funny. Alka-Seltzer is funny. Chicken is funny. Pickle is funny. All with a ‘k.’ ‘L’s are not funny. ‘M’s are not funny,” nor, might I suggest, are [ŋ]s.

Remember, I called Diller (and the Sterns) “innocent” language users. I can’t expect them to know any better. But linguists have no excuse. Even in the face of phonetic and cognitive experimental evidence to the contrary, why have the phonemists been so steadfast in their assertion that phonology is organized along phonemic lines? The answer, I suggest, is that phonemism is rooted in alphabetism. What then, is the route from alphabetism to phonemism? I turn to this question now.

Alphabetic writing has its origins in northern Semitic orthography, one which represents consonants, but not vowels. Between 2800 and 2700 years ago, the Greeks adapted and modified the Semitic system by introducing vowel symbols. This Greek innovation may seem like an obvious improvement over the vocally impoverished Semitic system (since all languages have both consonants and vowels) but its obviousness actually lies not in facts about phonetics, but rather, in facts about morphology. The morphology of Semitic languages (among them the present-day Amharic spoken in Ethiopia, the Arabic spoken on the Arabian peninsula and adjacent lands, and the Hebrew spoken in Israel) is that root morphemes may be analyzed as consisting of consonantal structure, while the inflectional system—tense, person, etc.—consists largely of vowel sounds that intervene between the root-based consonants. In Hebrew, for example, the three consonants ל [l], מ [m], and ד [d], when embedded in various contexts, all relate in some way to studying. Table 7.1 provides a few examples of both verbs and nouns. (Hebrew is written right-to-left.)

למדתי	l̥amadət ^h i	I studied	למדנו	l̥amadnu	we studied
למדת	l̥amadət ^h a	You studied (m)	למדסת	l̥amadət ^h εm	You studied (pl, m)
למדת	l̥amadet	You studied (f)	למדנת	l̥amadət ^h εn	You studied (pl, f)
למד	l̥amad	He studied	למדו	l̥amdu	They studied (m)
למדה	l̥amda	She studied	למדו	l̥amdu	They studied (f)
יילמד	jilmad	I will study	נלמד	nilmad	we will study
תלמד	t ^h ilmad	You will study (m)	תלמדו	t ^h ilmədu	You will study (pl, m)
תלמדי	t ^h ilmədi	You will study (f)	תלמדו	t ^h ilmədu	You will study (pl, f)
יילמד	jilmad	He will study	ילמדו	jilmədu	They will study (m)
תלמד	t ^h ilmad	She will study	יילמדו	jilmədu	They will study (f)
לומד	l̥omed	He studies	לומדים	l̥omdim	They study (m)
לומדת	l̥omədət	She studies	לומדתו	l̥omdot	They study (f)
למד	l̥omad	Study! (m)	למדו	limdu	Study! (pl, m)
למדי	limdi	Study! (f)	למדו	limdu	Study! (pl, f)
ללמוד	lilmod	To study	למידה	lemida	Studying (noun)
תלמיד	t ^h almid	Student (m)	תלמידה	t ^h almida	Student (f)
למודיה	limudija	Studio	למוד	limud	Study (noun)
תלמוד	t ^h almud	Book of (Jewish) law	למדן	lamdan	Researcher
למודי	limudi	Didactic	לומד	lumad	He was taught

Table 7.1. A partial list of Hebrew words with the root “למד”.

In isolation, many of these written words are what we might term “orthographically neutralized”; they are spelled identically. But in the context of written text, the intended word is quite clear to readers. Indeed, given the linear adjacency enjoyed by the root consonants in most forms, the specific root emerges quite effortlessly. Other consonantal roots, when embedded in these same contexts, produce words with the same inflectional properties, and are just as easy to read by anyone literate in Hebrew.

Due to the different morphological origins of consonants and vowels in Semitic, it becomes clear that the absence of vowels in the orthography is not a *shortcoming* of the system, but is instead an *innovation*: since vowel qualities are largely predictable by the grammatical context in which a given root is found, then representing them in the written code would be largely redundant. The northern Semites, then, devised a remarkably *efficient* orthographic system.

When the Greeks devised their own alphabet based on the Semitic model, they found the Semitic system wanting, because the vowel sounds of Greek morphology are not exclusively inflectional in origin, but instead are part of the root morphemes themselves. Consequently, the Greeks modified the nature of the Semitic system to include vowel symbols as well as consonant symbols; a useful innovation to be sure, but one that was naturally necessitated by the Greek morphological system. The Latin, Cyrillic, and Sanskrit alphabets—used for Indo-European languages with morphological systems broadly akin to Greek’s—all use vowel symbols as well.

Consider now orthographic systems that represent entire syllables, rather than individual consonants and vowels. These are known as *syllabaries*. Native Japanese orthography, for example, is syllable-based. Each symbol represents a particular combination of a vowel and a preceding consonant, if one; that is, an entire syllable. A Japanese syllabary is shown in Table 7.2.

あ a	い i	う u	え e	お o
か ka	き ki	く ku	け ke	こ ko
さ sa	し ji	す su	せ se	そ so
た ta	ち tji	つ tsu	て te	と to
な na	に ni	ぬ nu	ね ne	の no
は ha	ひ hi	ふ fu	へ he	ほ ho
ま ma	み mi	む mu	め me	も mo
や ja		ゆ ju		よ jo
ら ra	り ri	る ru	れ re	ろ ro
わ wa	ゐ wi		ゑ we	を wo
				ん ũ
が ga	ぎ gi	ぐ gu	げ ge	ご go
ざ za	じ ji	ず zu	ぜ ze	ぞ zo
だ da	ぢ ji	づ zu	で de	ど do
ば ba	び bi	ぶ bu	べ be	ぼ bo
ぱ pa	ぴ pi	ぷ pu	ぺ pe	ぽ po

Table 7.2. A native Japanese syllabary. The component parts of each syllable are not indicated, apart from the marker for obstruent voicing. See if you can find it.

Is a syllabary a less sophisticated orthographic system than an alphabetic one? Not at all; it is remarkably appropriate for the needs of its users. Japanese has a very limited number of syllable shapes. Consequently, a syllabary is a very efficient way to orthographically represent Japanese, although due to foreign influence—historically, Chinese, and more recently, American—Japanese has long used several different writing systems at once!

Finally, let's consider Chinese. As we discussed in Chapter One, each Chinese symbol, or character, represents a whole morpheme, and, like syllabaries, contains no information about the individual consonants and vowels that combine to form the phonetic quality of the morpheme. Surely, *this* is an inefficient system and inappropriate writing system, right? Well, yes and no. It is very inefficient in terms of promoting literacy, but really, it's a very natural development, given the uniquely centralist socio-cultural heritage of China. "Chinese" is actually a collection of five different languages—each of which possesses a myriad of dialects, many of which are mutually

unintelligible—that nonetheless are spoken by an ethnic group that has always been culturally, politically, and religiously homogenous. Although word-order is reasonably consistent across Chinese languages, the sound systems differ significantly from region to region. Consequently, a character-based writing system is the most appropriate means of making the written word accessible to speakers of all Chinese languages and dialects.

In general then, other than those all-too-common cases when orthography has been imposed from the outside by a religious or cultural imperial power, writing systems are appropriate to the linguistic and cultural system for which they have been developed.

So why has alphabetism been elevated to pre-eminence, among both lay people, and—especially for our purposes—as the type of writing system that (mis-) informs so many linguists’ intuitions about phonological structure? I suspect the answer has its origins in a further fact: there is a strong correlation between orthography and religion. Indeed, orthographic systems are often thought to be of divine origin. Jewish languages are written with Hebrew letters, the script used in the torah. The Kabbalists believe that mere exposure to the Hebrew letters on the written page may enhance one’s spiritual energy. Regarding the Devanagari script used for Sanskrit—the holy language of Hinduism—even its very name derives from the Sanskrit word for “divine”, *deva*. Arabic orthography—the system employed in the Koran—was brought to conquered lands when Islam was spread from the Arabian peninsula. Latin script—the orthographic system employed in the Christian scriptures—played a similar role in lands conquered by Christians. The Cyrillic alphabet—established by St. Cyril and St. Methodius, about 1300 years ago for Old Bulgarian—is used by Orthodox Christians. So whereas Orthodox Christian Serbs use Cyrillic script, Roman Catholic Croats—who speak the same language as the Serbs—use Latin script. In a very recent example, the Pahauh Hmong script was invented for the Hmong language of Southeast Asia in 1959 by Shong Lue Yang. Both Shong and his followers believed that he was a messenger from God to help the Hmong people.

This is surely not to say that modern-day phonemists believe in the divinity of the Christian-based Latin script—the script through which the worlds’ most influential linguists achieved literacy. Rather, these linguists are simply the inheritors of a system with an inherent Christocentric orientation, and so their wholly secular intuitions—intuitions that, for example, Classical Chinese linguists did *not* have—are nonetheless similarly tainted: the alphabetists of yesterday have become the phonemists of today.

As we discussed at length in Chapter Two, technological advances in the field of phonetics have demonstrated that the speech stream cannot be segmented into “letter”- or phoneme-sized chunks, and that the transitions between any steady-state components of the speech stream are the most informationally rich. Moreover, psychological experiments have shown that speakers who are not trained in an alphabetic orthography have no notion of the segment-sized chunks that you, I, and the phonemists take for granted. But even in the face of these uncontroversial findings, the phonemists have remained steadfast in their assertion that phonology is organized—at least at deep abstract levels—along phonemic lines. Why? Because they are relying on their *intuitions* about phonological structure. But these intuitions are directly traceable to their training in an alphabetic orthography, and so do not open any sort of window into the genuine structural properties of language. In short, the phonemists believe that alphabetic writing

is the most insightful orthographic system because it reveals the phonemic organization of phonological structure. But in fact, they get it exactly backwards: phonemists believe that phonological structure is built out of phonemes because they have been trained in an alphabetic orthography.

By embracing a mental entity for which there is no overt evidence—neither for the learner nor the linguist—the phonemists are forced to posit a multi-layered phonological system: the abstract level of the phoneme (the alphabetic level), the level of transformation (rules or constraints), and the output level (phonetic implementation), where the actual details that are present in speech are superimposed. As discussed in the previous sections, both “altruistic” and “selfish” approaches to phonology rely on the assumption there exists a speaker-controlled component that is in some sense “overlaid” on the more “authentic”, “underlying”, or “phonemic” phonological level.

So, when the phonemists observe the extra nasalization on a contrastively nasalized low vowel, they might attribute this to an altruistic linguistic level that is overlaid on the phonemic one, because, after all, the *phonemic* nasality of contrastively nasalized vowels is a pure, discrete, categorical phonological value, and hence is identical regardless of the vowel quality with which it is phonologically affiliated: at the phonological level, nasality is nasality is nasality. Consequently, when the phonemists observe different degrees of nasality on contrastively nasal vowels that differ in their oral qualities, (for example, more on [ã], less on other vowels), these differing degrees of nasality are “entirely unexpected”, and so must originate from another domain that is not purely phonological in origin.

And when they observe stops alternating with fricatives in intervocalic position, they might invoke “laziness” to account for their observations. According to the phonemists, despite their phonetic differences, the fricative and the stop are phonemically identical. The superficial phonetic fricative in Corsican (ɑ)[βok:ɑ] is an underlying phonological stop (/b/), based on linguistic “evidence” in the form of [boka].

When they observe the vowel nasalization that inevitably precedes a nasal consonant, they assign a different status to *this* nasalization as opposed to the nasalization that is co-extensive with the consonant itself. Recall that vowels have more nasality when an immediately following nasal consonant is made further back in the mouth, and have less nasality when an immediately following nasal is made further front in the mouth: the vowel in [dĩŋ] is more nasalized than the vowel in [dĩn], as discussed in Chapter Two. But because this vocalic nasality does not fit into to a letter- or phoneme-sized component of the speech stream (it is realized away from the nasal consonant itself, and might only partially overlap with a preceding vowel), these phonologists conclude that the nasalization is perhaps “merely” phonetic, or at least of a different phonological status than the nasal consonant, which is the “true” underlying phonological value. They propose this psychological distinction despite the fact that the vowel nasalization co-varies with the other cues to the oral component of the nasal consonant. That is, they propose this psychological distinction despite that fact that the vocalic nasality is functionally non-distinct from the nasal consonant itself; the complex is an integrated *Gestalt*.

When they observe the high component of rising tones being realized on a following vowel, as is found in Comaltepec Chinantec and many other languages (as discussed in Chapter Five), again, they often regard such effects as a mere phonetic

consequence: the underlying phonological affiliation—the “target”—of the high pitch is on the preceding vowel, but speakers are continually *missing* this target at the level of phonetic implementation. They are missing their target despite the fact that they have always produced their rising tones in exactly the way they have always *heard* them. So, according to the phonemists, speakers are continually *mentally undoing* aspects of the phonetic patterns that they hear, and then physically *re-instating* these patterns in their own speech! All this for the sake of maintaining the tenets of phonemism.

When they observe certain vowels in vowel harmony languages behaving in a *disharmonic* way—recall from Chapter Three that in Hungarian some roots with [i] take back vowel suffixes, because they historically derive from back vowels, for example, [hi:d] (“the bridge”) and [hi:dto:l] (“from the bridge”), with a back vowel suffix, was likely *hu:rd in Hungarian history ([u] is the back counterpart of [i])—the phonemists might treat this disharmonic *front* vowel as an underlying *back* vowel. In other words, Hungarian speakers today have the same abstract vowel system as did the Hungarian speakers of long ago. All that has changed is the phonological rule or constraint that transforms the underlying back vowel into a phonetic front vowel. Noam Chomsky and Morris Halle put it thus in 1968:

...[A]n observed [sound] change can only have one source—a change in the grammar that underlies the observed utterances.

A straightforward way of effecting changes in a grammar is to add new rules. The addition of a rule to the phonological component may be regarded as the most rudimentary type of sound change... Many sound changes known in diachronic phonology are of this type. By and large the familiar “sound laws” are, in fact, rules added to the phonological component...

...The conception of a linguistic change as a change in the grammar is also implicit in the traditional view of sound change. One of the crucial facts that linguists have tried to explain is that speakers are by and large unaware of the changes that their language is undergoing. The reason for this, it has been claimed, is that changes affect only the phonetic actualization of particular sounds—and, moreover, in so slight a measure that the changes appear to be gradual... Thus, vowels may be articulated somewhat farther back than before, or consonants may be actualized in some environment with aspiration of degree 4 whereas earlier they were actualized in that environment with degree 2. While there is no logical reason to reject this view of sound change, there is certainly no reason to give it special status. With the exception of the fact that speakers are unaware of an ongoing change—a fact which is easily explainable on the ground that speakers are, in general, unaware of the contents of their grammar—there is very little factual data to bear out this view... This lack of evidence, however, has not prevented scholars from continuing to espouse the gradual view of sound change.

The aforementioned are not just a few, isolated examples, by the way. Since Chomsky and Halle’s book of 1968, theorizing on the route between the proposed

underlying forms and actual surface forms (and back again) has pervaded the phonemists' phonological analyses; phonemism, far and away, is the dominant paradigm, although it should be said that rarely have phonologists gone so far as to invoke "altruism" or "laziness" in their attempts to defend this approach.

Sanford Schane, in his 1973 textbook on the Chomsky and Halle theory (called "generative phonology"), hints at some supposed "evidence" for phonemism, and circularly reasons the following:

Although the speech signal may be physically *continuous*, we seem to perceive it as a sequence of *discrete* entities.

That utterances can be represented as a sequence of discrete units is one of the basic theoretical assumptions of phonology... There is evidence that speakers also conceive of utterances as composed of discrete entities... [C]at is composed of three sounds. You might say that speakers are influenced by the written language, in which *cat* is spelled with three letters. But alphabetic writing can in fact be used as one argument in favor of the segmental view of speech, since with such writing systems there is a correlation between a sequence of graphic symbols and a purported sequence of speech sounds.

Larry M. Hyman, in his generative phonology textbook of two years later, asserts—as do Chomsky and Halle—that some phonetic components of the speech stream are redundant and predictable. According to Hyman, these supposedly redundant components of speech are not phonemic, but are instead merely phonetic. The phonemes themselves are not predictable, because any single combination of phonemes that makes up a morpheme or word is in some sense arbitrary; another combination of phonemes might have just as effectively been employed for a particular word. However, due to the particular combination of phonemes, certain supposedly *predictable* phonetic consequences may follow, albeit usually on a language-specific basis. For example, According to phonemists like Hyman, it is *arbitrary*—hence phonological—that the English word "pin" starts with a labial stop ("p"), but it is *predictable* that this labial stop will be aspirated in English ([p^h]). To reconcile the dichotomy between the proposed phonemes and the proposed redundancies, Hyman needs to further propose a multi-levelled phonology.

In order to characterize the relationship between the phonemes of a language and its inventory of phonetic segments, two levels of representation are distinguished, a *phonological level* and a *phonetic level*. Phonological representations consist of sequences of phonemes, transcribed between slashes (/.../); phonetic representations consist of sequences of phones, transcribed between square brackets ([...]). Thus the phonological representation of the word *pin* will be /pɪn/, while its phonetic representation will be [p^hɪn].

Since the phonological level represents the distinctive sound units of a language and not redundant phonetic information (such as the aspiration of the initial [p^h] of English /pɪn/), it is appropriate to think of it

as an approximation of the *mental* representation speakers have of the sounds of words in their language.

It appears to be immaterial to Hyman that both listeners and linguists are more likely to predict the presence of a [p] based on the particular formant transitions that take place during the [h] itself, rather than predict the aspiration based on the presence of the [p] closure. Clearly, the “phonemic” status of the “p” and the “non-phonemic” status of the [h] is a simple consequence of alphabetism—conveniently, we *spell* the word p-i-n, not p-h-i-n—and thus the distinction between redundant and non-redundant components of the speech stream—at least as characterized by Hyman—is revealed as a specious one. The criteria used to separate the redundant information from the phonemic content, then, is wholly arbitrary, and so, like Schane before him, Hyman lapses into circularity when trying to distinguish phonological from predictable content in the speech stream: he is driven by alphabetism.

Michael Kenstowicz and Charles Kisseberth draw a similar spurious distinction in *their* phonology text of 1979.

...[A]ll of the idiosyncratic features of the pronunciation of a morpheme are to be listed in the lexical representation of that morpheme in the lexicon of the grammar. On the other hand, features which are instances of systematic regularities will be assigned by phonological rules in the phonological component of the grammar. Correspondingly, we will make a distinction between two levels of representation of the phonological structure of a morpheme, word, phrase, or sentence: The *underlying representation* (UR), which will contain all the idiosyncratic information about the pronunciation of the constituent morphemes of the utterance, and the *phonetic representation* (PR), which contains the idiosyncratic information plus the predictable information about the pronunciation of the utterance.

Perhaps most boldly—in part because they state that a multi-leveled phonology is the *only* possible way to establish functional relationships among allomorphs—Carlos Gussenhoven and Haike Jacobs, in a textbook published in 1998, write:

Why do phonologists assume that there are two levels of representation, an underlying one and a surface one? Three arguments can be advanced.

1. One argument is economy. Why supply allophonic information in the lexical entries if it can be stated in a set of allophonic rules that are valid for all morphemes of the lexicon?...
2. [W]ith a single level it would not be possible to express the phonological relatedness of morpheme alternants...
3. Many generalizations are only valid at a level other than the surface level...

...The recognition of two levels of representation, a surface representation and a more underlying abstract representation, is the cornerstone of phonological theory.

Gussenhoven and Jacobs seem to be saying that phonologists *need* phonemes and the multi-leveled phonology that they demand in order to make their jobs easier, and they freely impart this *tool of the phonologist* a separate but equal *mental reality for the language user*.

Assertions like the four quoted above typically come in the first few pages of phonology texts, and in the first few lectures in courses on phonology. Although axiomatic in nature, they are presented to neophytes as the established *facts* on which all subsequent discussion is based. Indeed, the phonemists are quick to impose their (pre-) conception of phonological structure on their own language, and also on any other language that they analyze. And on those rare occasions when a pattern cannot be forced into the phonemic straightjacket, such as when its gradience or token-to-token variation is evident even to a phonetically untrained ear, the phonemists might regard the pattern as “a sound change in progress”, one that will ultimately be “phonologized” as one phonemic pattern or another. But of course, there is no such thing as “phonologization” in this sense of the terms meaning: at the proper level of description, *all* phonological patterns are sound changes in progress, as they are *all* gradiently and variably implemented, and they are *all* ever-changing. They are at once sufficiently stable to fulfill their communicative function, and sufficiently variable to be under constant modification: gradience and variation are the very stuff of phonology and sound change—listeners perceive it, and speakers produce it.

In sum, phonemism, which has its routes in alphabetism, requires phonologists to assume a multi-leveled phonology, an underlying level with the phonemes, and another level of derivation or transformation, in which all the supposedly predictable (non-phonemic; non-alphabetic) phonetic content is added to the underlying level. And because phonemes are, by axiom, genetically-endowed “building blocks” of linguistic structure, learners are pre-equipped with the linguistic faculties to figure all this out: despite the gradient and continuous nature of the speech stream, despite the inordinately complex way in which certain collections of phonetic properties might pattern together—co-vary—as a consequence of sound substitution (or might *not* pattern together due to an absence of alternation), learners can mentally undo all this complexity and arrive at a nice, neat system of a few stable elements that combine and re-combine. Then, as these listeners become speakers, they simply re-instate all the phonetic complexity that they have just mentally undone. All this for the sake of maintaining the tenets of phonemism.

Because of its derivational nature, phonemism typically involves a conflation of phonetics and cognition, and a conflation of diachrony and synchrony. Phonemism further forces a dichotomy between speakers’ supposed *mental intentions* and the actual *physical facts* of their speech: because the phonemists maintain the phonemic principle, they are forced into guessing games about speakers’ intentions. But indeed, the only intentions we may be confident that speakers possess are those that underlie speech acts themselves: speakers *intend* to speak, since they might just as readily choose to keep silent.

There are other ways the phoneme has been operationalized, however. In a landmark work published in 1939, the linguist Nikolai Trubetsky discussed a conception of the phoneme that comes much closer to the present approach. He defined the phoneme in purely functional terms, rejecting both the *phonetic* and *psychological* approaches that have long outlived Trubetsky himself:

Reference to psychology must be avoided in defining the phoneme since the latter is a linguistic and not a psychological concept. Any reference to “linguistic consciousness” must be avoided in defining the phoneme, “linguistic consciousness” being either a metaphorical designation of the system of language or a rather vague concept, which itself must be defined and possibly cannot even be defined... The phoneme is, above all, a functional concept that must be defined with respect to its function. Such a definition cannot be carried out with psycholinguistic notions... The phoneme can be defined satisfactorily neither on the basis of its psychological nature nor on the basis of its relation to the phonetic variants, but purely and solely on the basis of its function in the system of language... [E]very language presupposes distinctive (phonological) oppositions. The phoneme is a member of such an opposition that cannot be analyzed into still smaller distinctive (phonological) units. There is nothing to be changed in this quite clear and unequivocal definition. Any change can only lead to unnecessary complications.

For our purposes though, the phoneme is not an entity on any level—functional, phonetic, psychological, or even metaphorical. Rather, at best, “phoneme” is merely a terminological expedient that might capture the functional non-distinctness of any collection of phonetic properties that allomorphically alternates. And yet, despite its expedience, I choose to avoid the term altogether, because terminological expedients have a demonstrated tendency to become reified by their users.

LOQUOR ERGO ES

Let’s now reiterate some of the major points of all preceding discussion:

- Phonology is best characterized as a self-organizing and self-sustaining system of substantive social conventions which evolves passively over generations of speakers. It is self-organizing because its structural properties are a consequence of its use, requiring no outside monitor, guide, or force, to affect its organization. It is self-sustaining because, by its very use, it repairs and maintains itself.
- The regularities we observe in phonological systems are due to a complex interaction of phonetic and cognitive pressures acting over generations and generations of language use, and can be understood only when considering the communicative function of language itself.
- Present-day phonological patterns result from a dizzyingly complex and long-term interaction of articulatory, acoustic, auditory, aerodynamic, perceptual, functional, and social forces, and it’s facile to assume that the explanation for present-day phonetic and phonological patterns reduces to present-day factors.

- Our effortless mastery of sound substitutions derives from familiarity and experience with words themselves, and not from a mentally-compiled list of sound-sequential rules or constraints on what constitutes a “good” word.
- There is no reason to assume that language users subdivide the words they learn into distinct sound-components unless there is evidence from alternation to do so. Linguists may subdivide the speech signal into a myriad of constituents, but the only constituents that exist for language users are those which emerge as a consequence of language use, due to the linguistic functions that underlie them.
- The consonant-vowel sequences that we think we observe are simply artefactual, and it is the transitions between them that are most relevant, since they are the most informationally rich and often the most auditorily prominent components of the speech signal.
- It is impossible to account for the state of a language at a particular point in time without also considering the historical forces that have given rise to this state; sounds in alternation in the present often reflect sounds that have gradually changed in the past.
- Language is a system of conventionalized patterns of usage that arises from the minor and limited variations in which speakers naturally engage. The communicative success of certain spontaneous innovations over others—especially in the face of potentially confusing, homophonous forms—may very slowly, almost imperceptibly, drive the linguistic system in new directions.
- Very minor phonetic tendencies, coupled with the ambiguities they might induce or eschew, may eventually have far-reaching consequences for the phonological system.
- Our excellent-though-imperfect ability to engage in probability matching both causes and inhibits variation in speech. And the phonetic separation and stabilization of categories is as much a *consequence* of effective communication as it is a *cause* of effective communication.
- The presence of ambiguous tokens may result in listeners overestimating the prevalence of more distinct tokens. This overestimation, in turn, may result in more distinct tokens being produced, and, eventually, the better separation of phonological categories.
- Passive pressures toward homophone avoidance may trigger individual words to undergo pioneering changes.
- Phonological patterns evolve in the absence of the users who ultimately come to possess them.
- Present-day alternations have no present-day causes; they only have present-day effects.
- Alternations in the present—even when phonetically unnatural and superficially counter-functional—are the long-term product of small, local, and perfectly natural processes that play themselves out over generations of speakers.

- As phonetic pressures may influence a particular change in one sound, functional pressures may seize upon certain variants of another sound of the language, sending otherwise chance variation on a specific trajectory of change.
- Due to the inherent variability of speech production, and selectional pressures acting upon this variation, phonological systems evolve in fulfillment of their communicative function. It is the adaptation of a contrastive value to its context, and its subsequent survival as a functionally beneficial component of the communicative system, which is responsible for allophonic patterns.
- The origins of phonological simplifications lie not in speakers' supposedly unreliable production of speech, but instead in listeners' demonstrably reliable perception of meaning.
- Language use, language experience, and language function influence language structure, and any attempt to deny this fact will not advance our understanding of the organizing principles of language.
- Any approach to phonological structure which makes reference to speakers' supposed *intentions* should be regarded with suspicion. The only intentions we may be confident that speakers possess are those that are responsible for speech acts themselves: speakers *intend* to speak, since they might just as readily choose to keep silent.
- There is often evidence that one alternant *historically* derives from another, but there is no evidence to conclude that one alternant *psychologically* derives from another.
- All phonological patterns are sound changes in progress, as they are all gradually and variably implemented, and they are all ever-changing. They are at once sufficiently stable to fulfill their communicative function, and sufficiently variable to be under constant modification: gradience and variation are the very stuff of phonology and sound change—listeners perceive it, and speakers produce it.

Homo sapiens possesses a predisposition for language. This endowed trait emerged by adapting to (and evolving along with) traits that evolution had already provided to our evolutionary ancestors: the hearing mechanism, cognitive capabilities such as probability matching and pattern generalization, and the respiratory and ingestive systems—sound, mind, and body. But it is our predisposition toward socialization that truly sowed the seeds for this remarkable, quintessentially human adaptation: our species has evolved as a social one, and language is the culminating expression of this socialization.

As it is extremely unlikely that language is directly encoded in our genes, but is instead the emergent result of *other* genetically encoded traits, it is doubtful that language is directly susceptible to genetic mutation. Rather, language is a *fixed* and *immutable* characteristic of our species: if language were to change its fundamental characteristics, we would no longer be Homo sapiens.

Yet one of the most remarkable fixed and immutable properties of language is its built-in mechanism of *adaptation* and *change*. Due to the nature in which sound, mind, body interact, languages themselves *evolve*: the communicative success or failure of

certain spontaneous innovations over others may slowly drive the linguistic system in new directions.

Consequently, our fixed and immutable predisposition for language does not require us to posit fixed and immutable structural elements of language: we don't need to assume the pre-determined existence of phonemes or other phonological features, or a fixed system of phonological rules or constraints that we inherit as a consequence of our being *Homo sapiens*. We observe comparable patterns in language after language, and in era after era, not because of any fixed and immutable elements of the linguistic system, but because the same physical and cognitive laws and tendencies are continually acting and interacting in comparable ways. Collections of phonetic properties that *change* meaning, *eliminate* meaning distinctions, or *maintain* meaning distinctions might emerge and change as a consequence of their use, re-use, and disuse.

Indeed, virtually every characteristic of phonological systems that we have discussed in this book is amenable to *external* explanation—explainable in terms of systems that demonstrably serve *another* function. Only as a very last resort should we posit domain-specific content to language, or to any other natural system, for that matter: simple, general, and broadly applicable hypotheses are a hallmark of compelling scientific investigations. As of this writing, I have encountered no reason to breach the gates of this last resort in our attempts to explain phonological patterning. Both language itself, and its structural characteristics, are emergent consequences of our evolved status as social beings: *I speak, therefore you are*.

Mutatis mutandis, this quote from Charles Darwin's *The Origin of Species* serves as a fitting conclusion to all preceding discussion:

Although I am fully convinced of the truth of the views given in this volume under the form of an abstract, I by no means expect to convince experienced naturalists whose minds are stocked with a multitude of facts all viewed, during a long course of years, from a point of view directly opposite to mine. It is so easy to hide our ignorance under such expressions as the 'plan of creation' 'unity of design,' &c., and to think that we give an explanation when we only restate a fact. Anyone whose disposition leads him to attach more weight to unexplained difficulties than to the explanation of a certain number of facts will certainly reject my theory. A few naturalists, endowed with much flexibility of mind, and who have already begun to doubt on the immutability of species, may be influenced by this volume: but I look with confidence to the future, to young and rising naturalists, who will be able to view both sides of the question with impartiality. Whoever is led to believe that species are mutable will do good service by conscientiously expressing his conviction: for only thus can the load of prejudice by which this subject is overwhelmed be removed.

FURTHER READING:

On the Lombard Effect:

Lombard, Etienne (1911). “Le signe de l’élévation de la voix”, *Annales Des Maladies de l’Oreille et Du Larynx* 37:101–119.

Lane, Harlan, and Bernard Tranel. (1971). “The Lombard Sign and the role of hearing in speech”, *Journal of Speech and Hearing Research* 14:677–709.

The Paul quotes (“...[T]here is no such thing...”; “One of the commonest errors...”):

Paul, Hermann. (1890 [1970]). *Principles of the History of Language*. College Park: McGrath. Page 48; Pages 47-48.

The Schuchardt quote (“The change of a sound, ...”):

Schuchardt, Hugo. (1885 [1972]). “On sound laws: against the Neogrammarians”, in Theo Vennemann and Terence H Wilbur, *Schuchardt, The Neogrammarians and the Transformational Theory of Phonological Change*. Pages 57-58.

On altruism in phonology:

Kingston, John. (2002). “Keeping and losing contrasts”, Julie Larson and Mary Paster, editors, *Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society.

Kingston, John and Randy L. Diehl. (1994). “Phonetic knowledge”, *Language* 70:419-454.

On word frequency and sound change:

Bybee, Joan. (2001). *Phonology and Language Use*. Cambridge: Cambridge University Press.

Hale, Mark. (2003). “The ‘diachronic filter’ and phonological theory”, Paper presented at the 11th Manchester Conference on Phonology, Manchester, UK.

On laziness in phonology:

Kirchner, Robert, (2004). “Consonant Lenition”, in Bruce Hayes, Robert Kirchner, Donca Steriade, editors, *Phonetically Based Phonology*. Cambridge: Cambridge University Press. 313-345.

On writing systems:

Coulmas, Florian, (1996). *Encyclopedia of Writing Systems*. Oxford: Blackwell Publishers.

Smalley, William A., Chia Koua Vang, and Gnia Yee Yang. (1990). *Mother of Writing: The Origin and Development of a Hmong Messianic Script*. Chicago: The University of Chicago Press.

The Chomsky and Halle quote (“...[A]n observed [sound] change”):

Chomsky, Noam and Morris Halle. (1968). *The Sound Pattern of English*. New York: Harper and Row. Pages 249-250.

The Schane quote (“Although the speech signal may be physically continuous...”):
Schane, Sanford A. (1973). *Generative Phonology*. Englewood Cliffs, NJ: Prentiss-Hall. Pages 3-4.

The Hyman quote (“In order to characterize ...”):
Hyman, Larry M. (1975). *Phonology, Theory and Analysis*. New York: Holt, Rinehart, and Winston. Pages 8-9.

The Kenstowicz and Kisseberth quote (“...[A]ll of the idiosyncratic features ...”):
Kenstowicz, Michael, and Charles Kisseberth. (1979). *Generative Phonology*. San Diego: Academic Press. Page 32.

The Gussenhoven and Jacobs quote (“Why do phonologists assume...”; “The recognition of two levels of representation...”):
Gussenhoven, Carlos and Haike Jacobs. (1998). *Understanding Phonology*. London: Arnold. Pages 57 to 59; Page 62.

The Trubetskoy quote (“Reference to psychology...”):
Trubetskoy, Nikolai S. (1939 [1969]). *Principles of Phonology*. Berkeley/Los Angeles: University of California Press. Pages 38, 39, 41.

Other applications of evolutionary theory to linguistic theory:

Blevins, Juliette. (2004). *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge: Cambridge University Press.

Croft, William. (2000). *Explaining Language Change: An Evolutionary Approach*. Harlow, Essex: Longman.

Keller, Rudi. (1990 [1994]). *On Language Change: The Invisible Hand in Language*. London: Routledge.

Schleicher, August (1983 [1850]). *Die Sprachen Europas in Systematischer Übersicht: Linguistische Untersuchungen*. Philadelphia: John Benjamins.

The Darwin quote:

Darwin, Charles. (1859). *The Origin of Species by Means of Natural Selection*. London: John Murray.