Towards Progress in Theories of Language Sound Structure

Mark Liberman
University of Pennsylvania
myl@cis.upenn.edu

Towards Progress in Theories of Language Sound Structure

1. INTRODUCTION

The theme of this paper is that the traditional organization of phonological theories involves a crucial redundancy, and that serious consideration of this redundancy suggests a radical simplification of the theory. In technical terms, allophonic variation can be treated in two different ways: first, as a mapping from symbols to symbols, via phonological rules or constraints; or second, as a mapping from symbols to signals, via principles of phonetic realization. Careful examination of specific cases of allophonic variation generally suggests (and never seems to refute) a mode of description of the second type, in which structured phonological representations are mapped onto classes of phonetic trajectories. We should therefore consider the null hypothesis: a theory that entirely eliminates the symbolic treatment of allophonic variation, and makes post-lexical representations subject to direct phonetic interpretation, without any intervening symbol-manipulation, whether by rules or by constraints.

This leaves us with four well-motivated (indeed unavoidable) tools for dealing with sound-structure patterns:

1. Phonological inventory: The set of available phonological elements and structures.
2. Lexical entries: The phonological spelling of whatever entities are listed in the lexicon: roots, affixes, morphological templates, words, phrases.
3. Allomorphy: Alternative lexical pronunciations, whether conditioned by morphological features and morphological or phonological context, or in (linguistically) free variation.
4. Phonetic interpretation: the mapping between symbols (from 1-3 above) and signals.

And we assume, following Liberman & Pierrehumbert 1984 and many others, that patterns of phonetic interpretation are variably conditioned by structured phonological representations, as well as by rhetorical, attitudinal, and physiological variables, in ways that can be specific to particular language varieties and even particular individual speakers.

Given those resources, the phenomena generally described under the heading of *allophonic variation* do not require the addition of a fifth tool, in the form of manipulation of symbolic phonological representations via rules or constraints. Occam's Razor therefore suggests a null hypothesis that some may find surprising, since it is inconsistent with many aspects of the past century of phonological practice: Phonological rules or constraints of the traditional symbol-manipulating sort do not exist.

This is not a novel idea. For example, Liberman & Pierrehumbert 1984 suggested that

> Our methods … combine the phonologist's traditional concern for relations among abstract representations with the phonetician's interest in accounting for the details of actual speech. Our experience with these hybrid methods suggests that the correct "division of labor" between abstract phonological descriptions and the process of phonetic interpretation is not easy to discover. This point, applied to the subject matter of segmental phonology, will lead us to raise some pointed questions . . . about the correct treatment of allophonic variation. A reasonable answer to these questions would force most "postlexical" phenomena (in the sense of Mohanan (1982) and Kiparsky (1982)) to be treated as facts about the phonetic realization of phonological representations, rather than as modifications of phonological representations themselves. . . .
>
> In such a framework, the minimally required set of postlexical rules would combine lexical representations into a well-formed phrase-level phonological structure. One reasonable account of the division of labor between phonological representations and their phonetic implementation would limit postlexical rules to such a minimal set and assign all other postlexical regularities to phonetic implementation.

A version of this null hypothesis can be framed in nearly all of the many alternative theories of phonology, though it is easier to carry through in some theories rather than in others. The hypothesis may turn out to be false; but somewhat surprisingly, there do not seem to be any strong theoretical or empirical arguments against it at present. And recent technological

advances make it practical, for the first time, to mount careful empirical investigations of this null-hypothesis phonology as applied to a wide variety of relevant phenomena, as well as freeing us from the need to rely on arbitrary phonetic distinctions in order to make descriptive progress. Whatever the outcome, the search will enrich our understanding of language sound structure.

1.1 SYMBOLS AND SIGNALS

When someone speaks, they assemble a structured sequence of words, which they somehow encode as a pattern of vocal gestures and sounds. When all goes well, a hearer reverses the process and perceives the same structured word sequence. The individual elements of this word sequence are drawn from a set of entities that are crisply differentiated one from another, and transmitted from speaker to hearer with remarkable fidelity. In contrast, the corresponding vocal gestures and sounds are essentially continuous trajectories in continuous articulatory and acoustic spaces. Linguists traditionally model this situation by representing words as *symbols*, and vocal gestures and sounds as *signals*.

And as linguists for millennia have recognized, a second symbolic layer intervenes between words and sounds: *phonological representations*. These might be strings of phonemes as in traditional structuralist theories, or distinctive feature matrices as in Chomsky & Halle 1968, or linked autosegmental tiers as in Goldsmith 1976, or features arranged as moras, syllables and feet. But in any phonological system, a simple structured combination of a small finite set of basic symbolic elements defines the claims that words make on articulations and sounds.

We should note in passing that the history of automatic speech recognition technology underlines the motivations for (cultural or biological) evolution of a phonological system. We want accurate transmission of messages comprised of word-sequences drawn from a vocabulary of tens of thousands of items. But the signals representing these messages also carry many other sorts of information: about the identity of the speaker, about the acoustic environment, about the style and rhetorical structure of the message, about speaking rate and vocal effort and the speaker's

attitudes and emotions and so on. If words were arbitrary classes of vocal noises, learning a word would require hundreds or thousands of training examples from different speakers in different contexts, as was the case when ASR 'time warping' systems were based on stored word-level recordings. But when a word's pronunciation is encoded phonologically, every occurrence of every word helps us to learn the symbol-to-signal mapping of the phonological system, and therefore help us to learn to recognize all other words (in the same language) as well.

1.2 SOME OTHER RELEVANT CONCEPTS

There are many different theories about how to cash out these general ideas in formal or psychological detail. For the most part, these differences are orthogonal to the issue discussed here. Underlying all the alternatives is overwhelming evidence for the psychological reality and descriptive necessity of both lexical and phonological levels of representation as discrete, symbolic systems; and for the existence of variable, language-particular principles for the phonetic interpretation of phonological representations. However, we need to discuss a few additional concepts in order to clarify the proposal being made.

MORPHOLOGICAL STRUCTURE. There can be lexical structure inside wordforms (e.g. via inflection, derivation, and compounding) and also lexical entries that are phrasal in nature. These structures have consequences for pronunciation, through allomorphy and perhaps also through direct phonetic interpretation.

PHONOLOGICAL STRUCTURE. It's obvious that syllable structure, stress and foot structure, and phrasal structure play a role in speech patterns. For example, in most varieties of American English, the /t/ of 'at all' is pronounced as a voiced flap, whereas the /t/ of 'a tall' is pronounced as a voiceless aspirated stop. The relevant generalization is roughly that all non-onset consonants are weakened, and that in the case of intervocalic /t/ the closure is weakened to a ballistic tap and the laryngeal gesture is weakened to the point of disappearance. For phonetic interpretation to be an option in such cases, phonological structure must be available to be interpreted.

ALLOMORPHY. It's clear that lexical entries can have multiple phonological forms. In some cases these forms are in more or less free variation, like /ɛ/ versus /i/ in the first syllable of *economics*. In other cases, the choice of form depends on phonological or morphological context, like the two versions of the English indefinite article *a* and *an*. And via lexicalization, what starts as a casual-speech reduction can become an alternative lexical form, as in the case of English 'going to' in the version conventionally spelled 'gonna'.

EXEMPLAR THEORY AND WORD-SPECIFIC PHONETICS. Every theory has word-specific phonological representations (ignoring some radical and implausible proposals to do without phonology entirely). Some linguists (e.g. Bybee 2000, Hay 2000, Pierrehumbert 2001) have suggested that there might be lexically-specific phonetic interpretation as well. If true, this would add yet another descriptive option to an already over-complete set of alternatives, and the cited facts also lend themselves to accounts in terms of the well-established influence of frequency, register, dialect mixture, etc., on lexical choice and phonetic interpretation. But to the extent that lexically-specific phonetic interpretation exists, for example via lexical priming of gestural or acoustic variants, it further undermine the arguments for symbolic allophony.

QUASI-REGULARITY AND EMERGENT KNOWLEDGE. There are several strands of recent work (e.g. Rumelhart & McClelland 1986; Pierrehumbert 2001; Liberman 2004a, 2004b; Seidenberg & Plaut 2014) that blur the distinction between table lookup and derivation by rule, noting that it's possible to devise systems in which the learning of specific examples gradually generalizes to treat novel inputs in terms of similar patterns. As applied to lexical representations, this can be seen as a new form of traditional ideas about analogy. Patterns that look like symbolic allophony and allomorphy can emerge from such approaches to lexical storage. More radically, such ideas might blur the symbol/signal distinction entirely, or at least offer a story about how symbolic representations might emerge out of learned trajectories in signal space, while simultaneously generating those trajectories. At present, however, these ideas seem too amorphous and protean to define a productive approach to everyday linguistic description.

1.3 A LITTLE DISCIPLINARY HISTORY

There was a time when linguists were forced by circumstances to explore language sound structure almost entirely in symbolic terms. The shared perception of words anchored one end of the problem, and the principles that underlie alphabetic writing provided a rich and reliable array of discrete categories for characterizing the relationship of words to sounds, which were generalized to extend symbolic representations deep into the domain of signals.

Although these phonetic symbols often have rather poor intersubjective stability (see e.g. Shriberg & Lof 1991, Pitt et al. 2005, Oller & Ramsdell 2006), they are a convenient way to convey subjective impressions of contextual, dialectal, and historical patterns in language sound structure. And for more than a century, the manipulation of such symbols has been the standard method of describing such patterns. In the course of this process, linguists have evolved a number of ideas for generalizing representations beyond simple strings of symbols: phonological features, syllabic and prosodic structures, and the linked tiers of autosegments originally proposed by John Goldsmith.

In contrast to symbol-manipulation, quantitative measurements of speech signals, though clearly relevant, were once nearly impossible, and until recently were painfully difficult. And systematic models of such measurements, beyond ordinary statistical analysis, have been much less extensively explored. In recent years, technological innovation has changed this balance to some extent. Acoustic recording has become easy, and digital storage allows easy distribution of very large collections of sound recordings. Techniques borrowed from speech recognition give us convenient forced alignment of recorded audio to the words in transcripts, and increasingly reliable automatic classification and measurement of the pronunciation of those words. Database technology gives us instant access to arbitrary subsets of those annotated audio collections. We also have increasingly inexpensive and convenient articulatory measurement techniques such as ultrasound, electromagnetic articulography, and magnetic resonance imaging. And modern computer hardware and software makes it relatively easy to frame and test quantitative models relating linguistic symbols and signals.

This progress has a cost: linguistics faces an increasing embarrassment of theoretical and methodological riches. The new phonetic methodologies help us to discover many new patterns of sound in language. And for any set of such patterns, we can provide many descriptions and explanations that are conceptually very different, but describe almost exactly the same set of observations.

This was already an issue for purely symbolic accounts of linguistic sound structure, but a serious consideration of non-universal contextually-constrained symbol-signal mapping releases a host of new descriptive options. In particular, these new methods make it practical to frame and test, on a large scale, theories that dispense with symbolic allophony.

2. SOME ILLUSTRATIVE EXAMPLES

Across the world's languages, linguists have documented and analyzed thousands of cases of dozens of types of allophonic variation. Here we will take a brief look at just two well-documented phenomena. The point is not to prove that the proposed mode of analysis is correct, but simply to illustrate the sorts of patterns that arise, to sketch the way that linguists have treated them, and to suggest the alternative approaches that the proposed null-hypothesis phonology might take. If phonological theories without symbolic allophony should be treated as the null hypothesis, then we should accept the burden of proof to try to show that analyses based on such theories are inadequate.

2.1 CANADIAN RAISING.

Joos 1942 observed that in Ontario English,

> the diphthongs /aj/ and /aw/ … each have two varieties. One, which I shall call the HIGH diphthong after its initial tongue-position, begins with a lower-mid vowel-sound; it is used before any fortis consonant with zero juncture: [hwɛɪt, nɛɪf; ʃʌʊt, hʌʊt] = *white, knife; shout, house*. The other, the LOW diphthong, is used in all other contexts: [haɪ, faɪnd, naɪvz; hɑʊ, fɑʊnd, hɑʊzɪz] = *high, find , knives; how, found, houses*.

Joos described this phenomenon both qualitatively, in terms of his impressions of relative tongue position, and symbolically, in terms of the difference between [ɐ, ʌ] and [a, ɑ]. He does not describe it quantitatively, because in 1942 the quantitative measurement of vowel sounds was in practice impossible.

Joos suggests that the source of this difference was a "lesser movement of the tongue," associated with 'the relative shortness of English vowels before fortis consonants', which has been modified in the case of those two vowels 'from a difference essentially of length to a difference essentially of quality', conditional on a following 'fortis consonant with zero juncture', i.e. immediately following voiceless consonant within the same word.

Joos argues that this is enough to lay the groundwork for the contrast to become a matter of phonemic opposition (that is, change in lexical entries) vs. allophonic variation (that is, change by phonological rule):

> It is now possible for [ɐɪ, aɪ; ʌʊ, ɑʊ] to become four phonemes instead of two, while it would not have been possible if /aj, aw/ had continued to be split according to the same criterion as all other syllabics. This statement is here proposed as a criterion of the possibility of phonemicizing: A phoneme /x/ can be succeeded historically by two phonemes /x_1/ and /x_2/ only if there is a difference between the contrast [x_1] - [x_2] under the contextual opposition C_1 ~ C_2, and the contrast [y_1, z_1, ...] [y_2, z_2, ...] in other phonemes under the same C_1, C_2. Under this PRINCIPLE OF DIFFERENTIATION, when /k, g/ were split into Old English / č, ǰ / and /k, g/, there must have been a period when the articulatory contrast between [č, ǰ] before palatals and [k, g] elsewhere was different from the contrast between [p_1, b_1, t_1, ...] before palatals and [p_2, b_2, t_2, ...] elsewhere, but when [č, ǰ] and [k, g] were still allophones of /k, g/; for otherwise the palatalization-opposition would have affected all consonants equally, which is as much as to say that

> palatalization would have been a separate phoneme feature itself, or a classifier of part of the vowels: the history of palatalization in Russian is an example.

And he suggests that 'this possibility [is] beginning to become a reality', given the voicing of /t/ 'between voiced sounds with the syllable-division within it'. He asserts that in his own speech, 'it is not lenis [d]; it is a very short fortis', but 'in the speech of a large part of my contemporaries in the General American area, however, it has become a lenis [d], so that *latter = ladder* and *diluted = deluded*, with no difference in the preceding vowels either.' And he further claims that 'such speakers divide into two groups according to their pronunciation of words like *typewriter*. Group A says [tɛɪpaɪdə̴], while Group B says [tɛɪpɛɪdə̴].'

Joos also cites some lexical differentiation, noting that there are 'about two dozen common words like bite, biting' where Group A 'shifts /t/ to /d/ in the inflected forms while keeping the dipthong unchanged', whereas 'in hundreds of common words like *bet, betting*, there is also a difference in the vowels … so that *betting = bedding* in all its phonemes.' He asserts that 'This difference clearly establishes the phonemic splitting of the diphthongs', and suggests that 'From such a beginning, it need not be long before we hear high diphthongs before /b, g/ also, in contrast with low diphthongs, maybe in all homely words or on some such analogical basis'.

Without delving further into the interesting details of this paper, we can already see the crucial outline of the story. There are lexical representations, printed between slashes and consisting of strings what Joos calls 'phonemes', and there are pronunciations, printed between square brackets and consisting of strings of symbols from a similar set that Joos calls 'allophones'. Some phonetic changes are historically reanalyzed – here a length difference causes a 'lesser movement of the tongue' and is therefore reanalyzed in some cases as a difference in vowel quality. And some of these allophonic changes may become 'phonemicized' and thereby elevated to lexical status for some words. The interaction of allophonic patterns can create what has come to be called 'opacity', as when the voicing of intervocalic /t/ in certain inflected forms obscures the environment for Canadian Raising.

Joos subscribes to some very restrictive ideas about the nature and relationship of phonemes and allophones. And there are some factual questions about his treatment as well – his belief in his own voiced 'very short fortis' /t/ may well be an instance of the *phoneme restoration effect* (Warren 1970); and the existence of his Group B is controversial. The dozens of papers that have wrestled over the past 75 years with the issues that he raised have considered a much wider range of phonological theories, and have brought in a wider range of less anecdotal data, including the variants of /ay/ raising that have developed, apparently independently, in many other varieties of North American English.

But one long-outmoded aspect of Joos's treatment has stayed stubbornly in place in this literature. Although Joos recognizes the relevance of speech articulation and sound, through his mention of the 'lesser movement of the tongue', his description remains entirely symbolic, expressed in terms of relationships among symbol strings. He had no real choice in this matter, since in 1942 there was no accessible method for quantifying vowel quality. This changed in 1947 with the declassification of the sound spectrograph, and has changed further since that time with more and more accessible computer-based methods for measuring vowel formants or other proxies for vowel quality. And there have been a number of papers on /ay/ raising that take advantage of these methods and thereby shed additional light on the phenomenon, e.g. Thomas 2000, Moreton 2004, Moreton & Thomas 2007, Fruehwald 2007, Fruehwald 2013.

However, essentially all of the more recent treatments, whether or not based on phonetic measurements, continue to address the questions that Joos raised in essentially the same symbol-string terms that Joos used.

Thus Mielke et al. 2003 focus on ways to use the ranked constraints of Optimality Theory to show 'that some cases of opacity that were previously considered problematic for a surface-oriented formal model of synchronic phonology can be reanalyzed in a manner that renders the phonological patterns transparent', treating Canadian Raising as an alternation between the

phonetic strings [ay] and [ʌy]. They do note that

> While our analysis of Canadian Raising succeeds in transparently accounting for the observed data, we cannot help but speculate that a more satisfying explanation would directly incorporate the relationship between phonological voicing and preceding vowel length. … As Port (1996) observes, the tendency to view segments as discrete elements leads to the analysis of vowel raising and consonant voicing as separate contrasts, and this misses an important generalization about the interrelatedness of the phonetic realization of vowels and consonants in phonological contrast.

But this is no more than a modernized and expanded version of Joos's remark about a 'lesser movement of the tongue'.

Idsardi 2006 argues that 'recent efforts by Mielke et al. (2003) to revive Joos's 1942 phonemic splitting analysis … and to deny the existence of allophonic opacity are incorrect'. He mentions the 'growing industry in the phonetic measurement of the raised dipthongs', but argues that 'we need to resist the lure of the transcription systems', because

> The importance of Canadian Raising for opacity comes from its interaction with the process that neutralizes the /t-d/ contrast (or the neutralization of the /s-z/ contrast between $house_{Noun}$ and $house_{Verb}$ by phonetic devoicing), not from the phonetic details of the raising process itself.

In other words, the relevant issues arise within an essentially discrete, symbolic system, and we can remain agnostic about the distinctions' phonetic interpretation.

Pater 2014 offers yet another formal mechanism to account for 'the distribution of the raised variant of the Canadian English diphthongs'. In his system, 'the preflap raised diphthongs are licensed by a language-specific constraint … captured with a weighted constraint grammar'; and he 'shows how correct weights can be found with a simple, widely used learning algorithm'. He notes that

> As Idsardi (2006) points out, analyses of CANADIAN RAISING … are generally of two types:

> those that treat the low/raised diphthong distinction as phonemic (Joos 1942), and those that treat it as opaquely allophonic, with the surface vowel contrast derived from the underlying contrast between /t/ and /d/ that is itself neutralized to the flap.

Pater describes his proposal as 'a third type of analysis, intermediate between the phonemic and allophonic approaches, in which the distribution of these diphthongs is an instance of positionally restricted contrast'.

These intelligent and interesting analyses, along with many others that we could cite, share with Joos 1942 the property of seeing the problem in terms of the distribution of symbols on the phonological surface, in relation to their distribution in the basic lexical entries involved. And the interesting and significant research into the quantitative measures of vowel quality relevant to these phenomena does not fundamentally change this perspective.

Thus Moreton 2004 'found the /ai/ pattern of more peripheral F1 and F2 in the offglides /ɔi ei aʊ/ as well, showing that it is part of a general pattern of 'hyperarticulation before voiceless consonants'.. This claim about hyperarticulation essentially inverts Joos's 'lesser movement of the tongue' idea, and Moreton's measurements also show that Joos got the phonetic transcription wrong, since 'the diphthong nuclei were less affected than the offglides.' But this work echoes Joos in appealing to an articulatory and perceptual explanation of the forces leading to this sound change; and Moreton's careful production and perception experiments help to explain why similar changes have apparently developed independently in several different speech communities, without challenging the idea that these changes are symbolic.

Fruehwald 2013 tracks the /ay/ raising over several decades of sociolinguistic interviews from the Philadelphia Neighborhood Corpus, and produces striking evidence to

> [challenge] the conventional wisdom that phonologization is a late-stage reanalysis of phonetic coarticulatory and perceptual effects .... Rather, it appears that phonologization occurs simultaneously with the onset of phonetic changes.

He observes that

> the factors which categorize contexts as undergoing or not undergoing a change are best defined on phonological, not phonetic, grounds … Perhaps the most surprising result is that /ay/ raising has applied opaquely with respect to flapping from the very outset of its phonetic change. Despite the demonstrable phonetic differences between surface /t/ and /d/, and their flapped forms, /ay/ raising has always applied according to the underlying voicing of the following segment.

And he argues for a modified version of the "Big Bang" theory of sound change proposed by Janda and Joseph 2003:

> (6.4) The initial innovation in a conditioned sound change is phonological, thus abrupt.
>
> (6.5) The phonetic correlates of this abrupt phonological innovation are not necessarily large.

So Fruehwald uses sophisticated phonetic measurement and modeling to argue that 'the factors that characterize contexts' for sound change are phonological, i.e. in our terms 'symbolic' – and so is 'the initial innovation'.

As far as I can determine, none of the more than a hundred post-1942 treatments of this phenomenon gives serious consideration to the alternative account that would be forced by the null-hypothesis phonology under discussion here. In this alternative account, neither /ay/ raising nor the flapping and voicing of /t/ are symbolic changes, rewriting symbolically-expressed phonological representations. Rather, both /t/ and /ay/ are unchanged on the phonological surface, and are interpreted phonetically in diverse ways, depending on language variety and context, so as to reflect the observed patterns of pronunciation.

This account is consistent with overlaid processes of lexicalization, where some or all of the variation is moved up into the lexicon, by expanding the phonological inventory and/or modifying a suitable range of lexical entries. (And if we allow lexically-specific phonetic implementation, lexicalization does not require a change in the phonological inventory, at least

at first.) In fact, Fruehwald demonstrates that something of this kind has happened in Philadelphia, where some speakers have generalized the raised version of /ay/ vowel to a few words like *spider*.

Note that this account does away with all of the problems of opacity and 'counter-feeding order'. Contextually-varied phonetic interpretation of /ay/ and of /t/, treated independently, still results in the observed patterns, since (in the simplest case) each phonetic-interpretation pattern operates on an unchanging phonological representation. And given the options of modifying the phonological inventory and the phonological spelling out of individual lexical entries, the more complex outcomes remain easy to model.

2.2 Spanish /s/ lenition.

The weakening of syllable-final /s/ in Spanish has been even more widely studied than English /ay/ raising, but the intellectual histories of linguistic approaches to these two phenomena have been quite different. As discussed above, /ay/ raising has (nearly?) always been analyzed in terms of symbolic allophony, even when the analysis is based on instrumental phonetic measurements. In contrast, some analyses of Spanish syllable-final /s/ lenition have viewed it as an aspect of phonetic realization, while others have treated it in terms of the rewriting of a symbol string.

Thus Navarro 1948 offers considerable impressionistic detail about the variable realization of /s/ in various environments in Puerto Rican Spanish, always treating the process as a matter of phonetic detail rather than phonological change. For example:

> Delante de p, t, c (k), la aspiración de la s reduce y atenúa su sonido: *respeto, pestaña, pescar*. En las palabras de esta clase con *es-* inicial en principio de grupo, no solo la s sino toda la silaba se apaga hasta un grado casi imperceptible. … Sería exagerado decir que vocablos como *abispa, cresta, casco*, se convierten en *abippa, cretta, cacco*. En

> realidad se percibe siempre cierto resto de la aspiración entre la vocal acentuada y la oclusion siguiente.

And Lloyd 1987 places this approach in a broader historical perspective:

> The preponderance of open syllables in Spanish from earliest times, and the drive to make all syllables as open as possible, has had a continuous effect on syllable-final consonants. ...
>
> Related, in part at least, to the preceding phenomenon is the weakening of syllable-final /-s / into an aspiration which may eventually become so weak that it disappears.

On the other hand, much of the late 20th century literature treated this phenomenon in symbolic terms. Thus Beym 1963, describing Argentinian Spanish, discusses the geographic, social, and phonological conditioning of the symbolically-represented allophones [s], [z], [h], [ɦ], [s], [x], and [∅] (= zero) as variants of phonemic /s/ in syllable-final position. Ma and Herasimchuk 1972 limit their account of Puerto Rican Spanish /s/ to the three allophones [s], [h], and [∅]. Cedergren 1973 describes three 'relevant variants' of syllable-final /s/ in Panamanian Spanish, with several symbolically-repesented 'phonetic realizations' each:

Var.	Phonetic Real.
s	[s] [z] [sz]
h	[c] [h] [ɦ]
∅	[] ∅

Poplack 1980 writes that

> Puerto Rican Spanish (s) is variably subject to two weakening processes, aspiration and deletion, so that a phrase such as *las cosas bonitas*. 'the pretty things' can also be realized [lah 'kosah bo'nitah] or [Ia 'kosa bo'nita].

In some cases, these researchers assigned pronunciations to discrete classes in order to make the phenomena accessible to the then-popular computer programs for 'variable rules' in

sociolinguistics, which modeled the distribution of a binary variable using logistic regression. In other cases, the motivation is explicitly the problem of designing an intersubjectively valid coding scheme. For example, in a study of Cuban Spanish, Terrell 1979 writes:

> Certain methodological observations are in order. It was my intention to distinguish a variety of phonetic manifestations of /s/. However, it became quickly apparent that such a task, theoretically so simple, on a practical basis was impossible. It is imperative in any science to demand that others be able to replicate the results of any investigation. Replicability of the results would have been very difficult to achieve with a fine transcription. For this reason, the following system was selected.
>
> > s: all phones with some sibilance.
> > 0: complete absence of a phone representing /s/.
> > h: normally aspirated, sometimes very weak, often voiced or nasalized and possibly assimilated resulting in a geminate consonant cluster.
>
> It should be noted that this is essentially the same system used by others who have done quantitative studies of Spanish phonology.

In contrast, several more recent studies have used modern computer-based methods to solve Terrell's problem by applying appropriate regression models to the systematic measurement of spectral centroids, closure, voicing, and frication durations, etc., treating these measurements with appropriate regression models. Thus Fox 2006 writes:

> Automated speech recognition methods were used to code three dependent variables for a corpus of over 50,000 tokens of syllable-final /s/: deletion or retention of /s/, duration of retained /s/, and the spectral center of gravity of retained /s/. Multiple regression was performed for each of the dependent variables, on all of the data combined and on several subsets of the data.

Erker 2010 asserts that such approaches lead to greater insight:

> Among the most compelling and often replicated findings to emerge from socio-phonetic research is that correlations between linguistic form and social factors can be manifested not only at the level of the segment but also in fine-grained, subsegmental aspects of speech. That is, instrumental analysis has proven capable of uncovering systematic socio-phonetic variation within a single segmental category and also across more than one segmental category.

And File-Muriel and Brown 2011 reinforce this conclusion:

> Whereas previous studies of Spanish *s*-weakening have relied on impressionistic coding, the present study examines temporal and gradient acoustic details in the production of /s/ by eight females from Cali, Colombia, during sociolinguistic interviews. We propose a metric for quantifying *s*-realization by employing three scalar-dependent variables: *s*-duration, centroid, and voicelessness. The results of linear regressions indicate that the dependent variables are significantly conditioned by local speaking rate, word position, following and preceding phonological context, stress, and lexical frequency. This study sheds light on how each independent variable influences *s*-realization acoustically. For example, as local speaking rate increases, duration, centroid, and voicelessness decrease, which is indicative of lenition, and the same weakening tendency is observed when /s/ occurs in word-final position or is followed by a nonhigh vowel, whereas frequency contributes only to *s*-duration. We discuss the advantages of opting for instrumental measurements over symbolic representation.

As in the case of /ay/ raising, there is evidence for lexicalization. Thus Terrell 1979 writes:

> Informal experience and some direct work with the Spanish of illiterate and semi-illiterate Dominicans leads me to believe that many speakers in the Dominican Republic have speech with completely restructured lexicons, in which no word ends in /s/.

Bullock et al. 2014 argue that 'illiterate Dominicans are not "lost-s speakers" who arbitrarily add coda-s' as a form of hypercorrection, because

> If such speakers did exist, we should not find that they are able to adjust their rates of s-realization according to different conversational styles. Their linguistic performance should demonstrate an accidental or random realization of coda-s, where each token produced would as likely be intrusive as lexical. But we already have available evidence that this is not the case.

However, lexical identity does seem to be a relevant factor in determining rates of coda-s production in their data, suggesting that their speakers' lexical entries differ in the presence (or perhaps the strength) of s-less and s-ful variants.

To sum up, the literature on Spanish syllable-final /s/ lenition seems entirely consistent with the class of phonological theories suggested here, in which we rely entirely on the resources of the phonological inventory, the content of lexical entries including allomorphic variation, and contextually-varied patterns of phonetic interpretation, without any use of symbolically-defined allophony.

3. Conclusion

There at least two good practical reasons that scholars over the centuries have described allophonic variation in terms of the manipulation of phonological symbols.

One source for this practice is the description of historical change, where we see systematic correspondences in phonological representations across time and space. Since allophonic variation is often a form of change in progress, it's natural to treat it in the same way as we treat the raw materials of historical-comparative reconstruction.

And a second source is descriptive convenience: extensions of the usual inventory of phonemes, features, and phonological structures are an obvious way to keep track of impressionistic data about instances of pronunciation.

But historical change, by definition, involves changes in lexical entries and phonological inventories. Our proposed class of null-hypothesis theories includes ways for patterns of phonetic interpretation to be re-analyzed as changes in lexical entries and phonological inventories, and so the needs of historical-comparative reconstruction are not prima facie a reason to add symbolic allophony to our toolkit.

And modern methods of phonetic research allow us to extract and model quantitative acoustic measurements from thousands of hours of speech. If the phonetic phenomena are quantal, in the sense of Stevens 1972 and Stevens 1989, or otherwise fall into qualitatively different subsets, we can generate and apply appropriate (semi-) automatic classifiers. With increasing facility, we can select audio samples or generate artificial stimuli and run perception tests. So it may often remain convenient to use symbolic labels in discussing and thinking about allophonic variation, but we now have clear practical alternatives to the ontological commitment that this convenience too easily creates.

In sum, the symbolic treatment of allophony is a deeply-ingrained habit that our field should not continue to accept without evidence. Discarding this habit will give us a fresh perspective on familiar phenomena, and even if we end up persuading ourselves to take it up again, the experience will be instructive.

DEDICATION AND ACKNOWLEDGEMENTS: This essay is dedicated to John Goldsmith, from whom I learned the value of examining basic assumptions. Among the many people who have recently contributed to my thoughts on this topic, Larry Hyman and Neville Ryant deserve special acknowledgement. Errors and omissions are of course my own fault.

References

Bermúdez-Otero, Ricardo. 2014. Amphichronic explanation and the life cycle of phonological processes. The Oxford handbook of historical phonology. Oxford: Oxford University Press.

Bermúdez-Otero, Ricardo. 2014. Philadelphia /aɪ/-raising without rule insertion. Handout from Symposium on Historical Phonology.

Beym, Richard. 1963. Porteño /s/ and [h] [ɦ] [s] [x] [∅] as variants. Lingua 12: 119-204.

Bullock, Barbara, Almeida Toribio, and Mark Amengual. 2014. The status of *s* in Dominican Spanish. Lingua 143: 20-35.

Bybee, Joan. 2000. The phonology of the lexicon: evidence from lexical diffusion. In Usage-Based Models of Language, ed. by Barlow and Kemmer. Stanford: CSLI. 65-85.

Donegan, Patricia. 2002. Phonological processes and phonetic rules. In Future challenges for natural linguistics, ed. by Katarzyna Dziubalska-Kolaczyk & Jaroslaw Weckwerth. München: Lincom Europa. 57-82.

Cedergren, Henrietta. 1973. The interplay of social and linguistic factors in Panama. Doctoral dissertation, Cornell University.

Cedergren, Henrietta. 1986. Metrical Structure and Vowel Deletion in Montreal French. In Diversity and Diachrony, ed. by David Sankoff. Amsterdam: John Benjamins.

Chomsky, Noam & Morris Halle. 1968. The Sound Pattern of English. New York, NY: Harper & Row.

Erker, Daniel. 2010. A subsegmental approach to coda/s/weakening in Dominican Spanish. International Journal of the Sociology of Language 203: 9-26.

File-Muriel, Richard, and Earl Brown. 2011. The gradient nature of s-lenition in Caleño Spanish. Language Variation and Change 23(2): 223-243.

Fox, Michelle. 2006. Usage-based effects in Latin American Spanish syllable-final /s/ lenition. Doctoral dissertation, University of Pennsylvania.

Fruehwald, Josef. 2007. The Spread of Raising: Opacity, lexicalization and diffusion. NWAV 36.

Fruehwald, Josef. 2013. The phonological influence on phonetic change. Doctoral dissertation, University of Pennsylvania.

Goldsmith, John. 1976. Autosegmental phonology. Doctoral dissertation, Massachusetts Institute of Technology.

Goldsmith, John. 1990. Autosegmental and metrical phonology: A new synthesis. Oxford: Basil Blackwell.

Hay, Jen. 2000. Causes and Consequences of Word Structure. Doctoral dissertation, Northwestern University.

Idsardi, William. 2006. Canadian raising, opacity, and rephonemicization. The Canadian Journal of Linguistics/La revue canadienne de linguistique. 51(2). 119-126.

Janda, Richard, and Brian Joseph. 2003. Reconsidering the Canons of Sound-Change: Towards a 'Big Bang' Theory. In Historical Linguistics 2001. Selected Papers from the 15th International Conference on Historical Linguistics, ed. by Barry Blake and Kate Burridge, 205-219. John Benjamins.

Joos, Martin. 1942. A phonological dilemma in Canadian English. Language 18(2). 141-144.

Liberman, Mark & Alan Prince. 1977. On Stress and Linguistic Rhythm. Linguistic Inquiry 8(2): 249-336.

Liberman, Mark, and Janet Pierrehumbert. 1984. Intonational invariance under changes in pitch range and length. In Language sound structure, ed. by Mark Aronoff and Richard Oehrle, 157-233. Cambridge, MA: MIT Press.

Liberman, Mark. 2004a. The Theology of Phonology. Language Log 1/2/2004.

Liberman, Mark. 2004b. The Curious Case of Quasiregularity. Language Log 1/15/2004.

Lloyd, Paul. 1987. From Latin to Spanish. Philadelphia: American Philosophical Society.

Ma, Roxana, and Elenor Herasimchuk. 1972. The Linguistic Dimensions of a Bilingual Neighborhood. In Bilingualism in the Barrio, ed. by Fishman, Cooper and Ma. Bloomington IN: Indiana University Press.

Mielke, Jeff, Mike Armstrong, and Elizabeth Hume. 2003. Looking through opacity. Theoretical Linguistics 29.123–139.

Moreton, Elliott. 2004. Realization of the English postvocalic [voice] contrast in F1 and F2. Journal of Phonetics 32(1): 1-33.

Moreton, Elliott, and Erik R. Thomas. 2007. Origins of Canadian Raising in voiceless-coda effects: a case study in phonologization. Laboratory Phonology 9.37-64.

Navarro, Tomas. 1948. El Espanol en Puerto Rico, Contribucion a la Geografia Linguistica Hispanoamerica. Río Piedra P.R.: Editorial de la Universidad de Puerto Rico.

Oller, Kimbrough, and Heather Ramsdell. 2006. A weighted reliability measure for phonetic transcription. Journal of Speech, Language, and Hearing Research. 49)6_: 1391-1411.

Pater, Joe. 2014. Canadian raising with language-specific weighted constraints. Language 90(1).230-240.

Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In Frequency effects and the emergence of linguistic structure, ed. by Bybee & Hopper, 137-158. Amsterdam: John Benjamins.

Pitt, Mark, Keith Johnson, Elizabeth Hume, Scott Kiesling, and William Raymond. 2005. The Buckeye corpus of conversational speech: labeling conventions and a test of transcriber reliability. Speech Communication 45(1): 89-95.

Poplack, Shana. 1979. Function and Process in a Variable Phonology. Doctoral Dissertation, University of Pennsylvania.

Poplack, Shana. 1980. The notion of the plural in Puerto Rican Spanish: Competing constraints on (s) deletion. In Locating language in time and space, ed. by William Labov, 55-67. New York: Academic Press.

Rumelhart, David, and James McClelland. 1986. On learning the past tenses of English verbs. In Parallel Distributed Processing: Explorations in the microstructure of cognition, ed. by McClelland and Rumelhart. Cambridge MA: MIT Press.

Seidenberg, Mark, and David Plaut. 2014. Quasiregularity and its Discontents: The Legacy of the Past Tense Debate. Cognitive Science 36(6): 1190-1228.

Shriberg, Lawrence, and Gregory Lof. 1991. Reliability studies in broad and narrow phonetic transcription. Clinical Linguistics & Phonetics 5(3): 225-279.

Stevens, Kenneth. 1972. The quantal nature of speech: Evidence from articulatory-acoustic data. In Human communication: A unified view, ed. by Peter Denes and Edward David Jr., 51–66. New York NY: McGraw Hill.

Stevens, Kenneth. 1989. On the quantal nature of speech. Journal of Phonetics, 17.3-45.

Stevens, Kenneth, and Keyser, Samuel. 2010. Quantal theory, enhancement and overlap. Journal of Phonetics 38(1).10-19.

Terrell, Tracy. Final /s/ in Cuban Spanish. Hispania 1979:599-612.

Thomas, Erik. 2000. Spectral differences in /ai/ offsets conditioned by voicing of the following consonant. Journal of Phonetics 28(1): 1-25.

Warren, Richard. 1970. Perceptual Restoration of Missing Speech Sounds. Science 167(3917): 392-393.