An I-Language Approach to Phonologization and Lexification

Mark Hale, Madelyn Kissock & Charles Reiss

1. Introduction

A very real problem with adopting the I-language perspective (e.g., Chomsky 1986) in diachronic linguistics, is that doing so makes it difficult to talk about even the most well-established results and the simplest notions. A statement like "The English word *knight* began with a *kn* cluster in Middle English" seems innocuous enough, until we realize that there is no scientifically useful or coherent definition of "English", or "Middle English", or "the word *knight*"! There is obviously no direct sense in which a mental representation in Chaucer's mind/brain (or even a sound Chaucer made) that we might refer to as "Chaucer's word for *knight*" has turned into a representation or sound that we might refer to as "Chomsky's word for *knight*".

One strategy for dealing with such difficulties is to ignore them—it is shockingly easy, as we ourselves have done for much of our professional lives, to teach students and write papers as if English, Marshallese, and Vedic Sanskrit were legitimate objects of scientific study. Such work engages our intellect and provides a certain amount of satisfaction, but our acceptance of the correctness of the I-language approach, and the problems and inconsistencies that arise by not adopting that approach, sometimes force us to confront the contradictions that we ourselves have helped to perpetuate. There is no question that doing so is difficult and sometimes tedious, but we hope the investment will generate a deeper understanding of what is traditionally called "language change", and ultimately allow us to reformulate informal statements, like the one about "the word *knight*", in terms that do not conflict with our basic scientific understanding of languages as properties of individuals.

There is a tradition among generative linguists interested in diachrony of treating "sound change" or "phonological change" as due to direct adult modification of phonological knowledge—sound change is a change in a speaker's phonology (see discussion in Hale 2007, chapters 3 and 6). By explaining sound change as a change in a speaker's grammar, his/her I-language, one can explain the regularity of sound change, the Neogrammarian Hypothesis, thus: since grammars are computational systems, a change in a speaker's grammar will generate regularities in output *vis-à-vis* the output of the pre-change stage. We have argued (Hale 2003, 2007), however, that so-called language "change" is *not* due to

spontaneous adult modification, so we are left without an explanation for regularity. In this chapter, we review arguments against viewing change as occurring in adult phonologies and arguments for localizing "change" in the imperfect transmission of grammar to learners. In other words, in cases of "change" there is no entity that changes, no change of the grammar of a speaker or the output of a speaker. Rather so-called "change" is a relationship between the grammars or the outputs of the grammars of speakers of different generations. Since this process does not involve phonological computations from underlying forms to surface forms in a single grammar, there is, strictly speaking, no phonology involved in (this kind of) sound change. We thus arrive at the position that phonological change is neither 'phonological' nor 'change'. We offer a partial solution to the puzzle of the regularity of sound change under this view.

The foundation of our approach was laid out most fully in Hale (2007). In this chapter, we review and develop some of those arguments, but our primary purpose is to apply the framework to the study of a particular phenomenon known as *phonologization*, and a putatively related phenomenon called *phonemicization*.

2. Phonologization

Phonologization is a label for a variety of diachronic processes in which some observable articulatory or acoustic phenomenon which arises in the course of speech without being explicitly represented at the grammatical level (e.g., a coarticulation effect) comes to be represented in the phonology. Though of course by no means the first work to recognize the concept, the *locus classicus* for modern discussions of the phenomenon is clearly Hyman (1976). That paper uses the example of the lowering of pitch on a following vowel intrinsic to the articulation of voiced stops, distinguishing three stages in the possible diachronic effects of that lowering, shown in Figure 1.

Stage 1	Stage 2	Stage 3
pá [˥]	pá [˥]	pá [˥]
bá [˧˥]	bă [˩˧]	pă [˩˧]
pitch lowering as a result of articulatory processes	(exaggerated) lowering of pitch as a result of phonological computation	now redundant voicing eliminated on [b], resulting in phonemic tone contrast
	phonologization	phonemicization

Figure 1: Three stages in the relationship between stop voicing and pitch

At Stage 1 there is a lowering of the pitch on the left-edge of a high-tone vowel due to the articulation of voiced stops.[1] At Stage 2, this lowering has come to be introduced by phonological rule, rather than being a mechanical by-product of articulation.[2] As such, it is realized in what Hyman calls an 'exaggerated' fashion (i.e., there is greater pitch lowering). Finally, at Stage 3 the now redundant (because of the induced tonal effects) voicing on stops is eliminated, leaving an unconditioned (i.e., phonemic) tonal contrast.

Hyman labels the diachronic development from Stage 1 to Stage 2 'phonologization,' reserving the term 'phonemicization' for the development from Stage 2 to Stage 3.[3] While this usage is still not completely standard,[4] it does seem that most technical discussions of 'phonologization' in the current literature maintain Hyman's distinction, and we will do so here. Sitting as it does at the intersection of both the complex issues surrounding the interface between phonetics, broadly construed, and phonology, and the equally problematic matters which arise when one seriously considers the relationship between the synchrony and the diachrony of linguistic systems, it will come as no surprise that 'phonologization' itself is a rich and complex topic for diachronic phonology.

Rather extensive terminological clarification will be required before we turn to a discussion of this concept. The goal of this clarification is not to advocate one particular terminological system over another, but rather to insure conceptual transparency in the discussion which follows. The reader must bear in mind that the terminology we will use here will not, or rather would only accidentally, be in agreement with that found in the literature on phonologization, which is itself internally inconsistent, and thus cannot form the basis for coherent discussion.

3. Terminological Matters: Synchronic

[1] See Kingston's (to appear) treatment of tonogenesis for exploration of the phonetic motivations of such interactions of voicing and pitch/tone.

[2] Although the term 'rule' is most appropriate when discussing Hyman's analysis, the points made in this paper are not tied to a particular theoretical framework beyond one that has, as its fundamental assumption, the notion of the grammar as a computational system. We therefore use 'rule', 'operation', and 'computation' indiscriminately in what follows.

[3] Hyman's discussion is compatible with our view that each Stage corresponds to a different grammar (in different individuals)—there is no reason to think that a person's grammar transitions from one Stage to another, but see discussion below.

[4] For example, the *Oxford Concise Dictionary of Linguistics* defines 'phonologization' as the '[h]istorical process by which a phonetic difference becomes a difference between phonemes'—i.e., as being equivalent to Hyman's 'phonemicization'.

If 'phonologization' is to mean anything at all, we must, in order to discover the properties of phonologization events, draw a clear and distinct division between the domain of phonetics (again, very broadly construed) and the domain of phonology—'phonologization' is, after all, *about* crossing that line.[5] As if the difficulties surrounding the efforts to establish a meaningful line were not themselves of sufficient complexity, the discussion of this matter has generally been plagued by an unfortunate ambiguity as to the meaning of 'phonetics', arising at least in part from the use of the symbols of the International Phonetic Alphabet to represent objects of radically different types. A conceptually straightforward division between the physical ('phonetics') and the mental ('phonology'), which is probably the closest approximation to historical conceptions of the place of the division (at least in the early days of the concept of the phoneme) is, in the end, not adequate. Nor would any conception of these matters which fails to consider the physical, limiting itself to the mental, or the mental, limiting itself to the physical, be able to provide a meaningful foundation for a theory of 'phonologization.' We need both the mental and the physical, and it seems clear that the physical facts (articulatory, acoustic, or aerodynamic) must fall outside the domain of phonology proper. It is, however, just as clearly inadequate to adopt the position that all mental aspects of speech generation are necessarily 'phonology' in the sense we need[6]

A sensible way of dividing things up must recognize that there is a great deal

[5]There appears to be great resistance to this move in the 'phonologization' literature. Hyman (2008), which provides an overview of his work in this area over many years, states as one of his conclusions that "there is overlap and unclarity as to where phonetics ends and phonology begins", as if this were an ontological problem with the world, instead of reflecting terminological inconsistency and our collective ignorance as a field. The way to establish the most productive location for a terminological division between phonetics and phonology is to assert some well-defined and coherent division, and explore the implications of that division for the models of the world we are constructing. None of us knows where a useful line is, we may never know, but we will never find the line if we fail to explore the implications of drawing it in some particular place.

[6]We have found it useful (Hale, Kissock & Reiss, 2007) to compare two excellent articles, Keating (1988) and Hammarberg (1976) which both attempt to define a usable phonetics-phonology boundary. For Keating, a phenomenon is "phonological" if it involves features. For Hammarberg, a phenomenon is "phonological" if it involves cognition, whether reliant on something like traditional discrete features or non-featural motor planning. Both definitions are coherent, and the cogency of the two positions suggests a need for a three-way contrast in terminology. However, for present purposes, we follow Keating's two-way distinction, collapsing cognitive and non-cognitive "phonetics" in the following discussion (but distinguishing these from the so-called 'phonetic output' level of phonological computation).

which is mental but not strictly speaking within the computational system which we call 'the grammar' (which system itself houses 'the phonology').[7]

In Figure 2, line (A) is the input to the phonological computation. Such an input contains the long-term stored form of a particular morpheme, or a combination of morphemes (say, a root with affixes, all stored in the lexicon). This input to the phonology is sometimes called the 'Phonemic' or 'Phonological' or 'Underlying' Representation, and it is traditionally placed within slash brackets. We assume that these representations are made up of 'features'.[8] Obviously, many different conceptions of the specific features which distinguish underlying representations are consistent with this understanding of things.

We have placed 'phonemic' in scare quotes above because, although this is often glossed over in modern phonology, these representations are not 'classical' phonemes in the, e.g., American structuralist sense. Nor are they 'classical' archiphonemes, in the European structuralist sense. What they really are is a little unclear at present, giving rise to a great deal of difficulty (discussion of which is generally absent) about just how abstract these representations may be, and what empirical facts must be observed to justify some particular degree of abstractness in any specific case. These concerns lie outside the coverage of this chapter. In *SPE* (Chomsky & Halle, 1968:11) the term 'phonemic' is avoided, including 'systematic phonemic' and 'morphophonemic'. Instead, both lexical representations and the phonological representations that are the result of syntactic Surface Structure and readjustment rules are referred to as "underlying representations", in other words, the input to phonology. We follow this practice here, since the difference between single and multiple morpheme inputs is not relevant to our point.

Line (B) of Figure 2 corresponds to the computational system which, in keeping with common practice, we call the 'phonology'. This system is a function

[7] It must be borne in mind that, as will become apparent in the subsequent discussion, the IPA characters in use in Figure 2 represent feature bundles at the 'phonemic' and 'phonetic output' levels, but impressionistic acoustic/articulatory representations at the level of bodily output. It is the ambiguous use of IPA symbols in these two very distinct functions that was alluded to earlier.

[8] This would seem to follow by virtual conceptual necessity: the features designate (abstract) properties of these mental objects—how could two representations be distinct without that distinction being due to some difference in properties? The features are the properties that make representations different from one another, and it is these properties which are subjected to modification in the course of phonological computation. Again, since there are no *other* properties of the segments than those designated by their features, phonological computation has nothing else to manipulate.

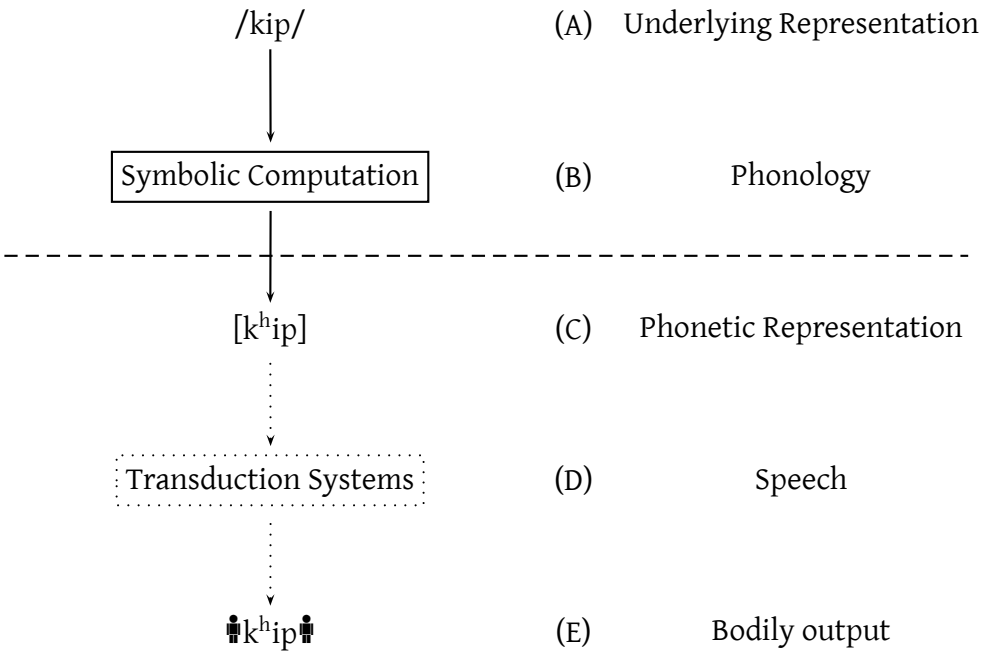

Figure 2: Speech Chain: Phonology, Phonetics, and Speech

mapping phonological properties of the input, (A) to a Phonetic Representation (C). It appears fairly clear that this computational system produces outputs in the same 'representational alphabet' as that provided by UG for the representation of long-term stored forms of the type that constitute (A):[9] feature bundles, syllables, feet, moraic structure, and the like are some possible elements of this representational system. Obviously, the computational system may be any of a wide range of possible types (including both rule- and constraint-based systems)—the question remains open.

Line (C) is referred to variously as the output (of the phonology)[10], the Surface

[9]Optimality-theoretic notions such as Faith and other input-output correspondence relations would seem to require such an identity: How could input-output identity ever be satisfied if inputs and outputs consisted of disjoint sets of properties?

[10]Even a term like "output" is sometimes ambiguous. Within the OT tradition the most harmonic candidate is the output of the grammar, but each candidate is referred to as a (potential)

representation, and the Phonetic Representation – we adopt the latter term, following *SPE*. We follow the tradition of placing this type of representation within square brackets. Crucially, it does not seem that the output representation (C) requires anything beyond the representational capacities used for underlying representations; i.e., there is no 'phonetic representation' system (in this sense of 'phonetic') distinct from the phonological one.

This identity in representational systems allows us to classify the phonology as involving *computation*, rather than *transduction*, which involves a conversion of the nature of the objects involved[11] The conversion of the phonetic representation (C) to actual bodily output (E) (i.e., sound, in the case of speech) is clearly transduction (see below), and the phonology is just as clearly computation. In Figure 2 we have contrasted transduction (dotted arrows) with computation (solid arrows).

Thus, as noted above, 'the phonetic representation' (C) is a mental object, unaffected by factors like speech rate, muscular fatigue and mucosal lubrication, generated using the same representational system as that provided by UG for the long-term storage of phonological representations. Its status as different from the underlying representation in (A) arises thus not from any particular properties that it has, but rather by virtue of the place in the system at which it comes into existence—it is a short-term computational output in the phonological alphabet. It thus contrasts with the underlying form (A) which contains the representations stored in the lexicon (see note 7); and in its symbolic nature and representational alphabet, it contrasts with the bodily output (E).

A set of poorly understood and relatively complex systems intervene between the phonetic output representation (C) and the actual bodily output (E), which is denoted by placement within 'body' brackets, a practice that we hope will become standard in the field, following Hale 2007. These systems include both factors under some cognitive control (the generation of a 'gestural score', speech planning, rate and loudness considerations, etc.) and more 'chaotic' ones outside of cognitive control (current physical properties of the vocal tract, air pressure, etc.). Some of these systems may be 'computational' in nature, others 'transducers', but when lumped together as space (and competence) limitations compel us to do here on line (D), the entire set of systems taken together act as a transducer, since we get a mapping between phonetic representations and the

output as well, for example, when evaluating Output-Output Identity constraints.

[11] See Hale & Reiss (2008:109) for discussion of this term. The air pressure differentials converted to patterns of electron pulses by a microphone would be a good example of transduction, since the elements involved are of radically different types in the two cases.

articulatory gestures that initiate sound waves.

This transduction of the phonetic representation into some bodily realization is, under this conception of things, extra-grammatical, involving both non-grammatical mental computation (e.g., motor planning) and non-grammatical physical modification of the implementation of the motor plan (e.g., salivary interference). For convenience, we call this the 'speech planning and implementation', or, for short, the 'speech' level.

Where in such a figure are we to draw the line which will form the central concern of a theory of 'phonologization'? Here the literature shows some variety of opinion, but, also, rather too frequently, evidence of no opinion at all—i.e., great inexplicitness and confusion. IPA-type transcriptions are used to describe both the output of phonological computation (our 'phonetic representation'—the (C) of Figure 2) and the impressionistic rendering by a researcher of the 'phonetic' bodily output of a speaker (E).[12] Claims are then made about 'phonetic' transcriptions, some of which are true of *only* those that correspond to line (C), others that *only* in the latter role correspond to (E): these claims are then conflated, leading to the aforementioned confusion. A cursory examination of the use of the word 'allophone' in almost any introductory textbook on phonology will clearly reveal the confusion at play.[13]

Wherever the line is to be most productively and insightfully drawn, it is clear that the very worst move in trying to discuss technical matters such as 'phonologization' is to fail to draw the line sharply and clearly; from a vague and confused line will come only vague and confused conceptions of 'phonologization'. A very clear place to draw the line, consistent with some past uses of the terms in question, is between (B) and (C) of Figure 2. The (combinations of) long-term stored phonological representations, the underlying representations in (A), are the input to the phonology (B), the computational device which modifies those representations. The 'phonetic representation' (C) is the output of the phonology. In our view, a coherent definition of phonetics is the mapping from (C) to (E), via the many complex systems of (D). We therefore place the phonetics-phonology division just above (C).[14]

[12]We have continued this confusing use of IPA symbols, but use our 'body' brackets to keep the reader alert to our transgression.

[13]There is a small, mostly ignored literature on the distinction between intrinsic and extrinsic allophones that tries to tease these issues apart (e.g. Tatham 1971).

[14]Our model of the speech chain in Figure 2 contrasts sharply with the widespread advocacy of what is sometimes called 'language-specific phonetics.' The literature on the matter is extensive (e.g., Kingston & Diehl 1994, Kingston 2007 with references therein) , and we have addressed it

4. Terminological Matters: Diachronic

Because 'phonologization' is a type of language change, we must also clarify certain diachronic issues. The fundamental matter, from which further clarity can be made to follow, is this: change is not a property of grammars, but rather a relationship between grammars. We refer to G_1, which existed before the change, and G_2, which exists after the change, in an attempt to interpret Hyman's Stages in I-language terms.

Adopting the I-language approach, we see that in spite of the strong superficial similarity between the synchronic and diachronic versions of a statement such as '*k* becomes *ʔ* word-finally', the processes underlying these two kinds of events are radically distinct. A synchronic modification of the feature bundle /k/ to the feature bundle [ʔ] in the phonology represents computation within a human mind, necessarily constrained by the principles of UG which give rise to phonologies in humans. The diachronic version of this modification represents a situation in which one grammar, G_1, manifests 🧍k🧍-type outputs, while G_2, crucially a 'descendant' of G_1 (a definitional matter to which we will return shortly), manifests 🧍ʔ🧍-type outputs in *corresponding* positions in *corresponding* lexical items.

The crucial concepts in the diachronic domain under this interpretation of change are that of the 'descent' relationship between distinct grammars and that of the 'corresponding' lexical item (and 'corresponding' segment within corresponding lexical items). Neither of these concepts is simple, and neither is particularly clearly articulated in the literature on diachronic phonology. The latter of these can probably be safely left at the intuitive level here (as is the norm in both diachronic and synchronic phonology)—the complications are well-known, if still not fully under our conceptual control.[15]

As to the former matter—the question of when two grammars can be said to stand in a relationship such that it is appropriate to refer to G_2 as having arisen via 'change' from G_1—the only clear answer that has been offered to this question is the following. G_2 arises from G_1 via change only if G_1 provided the

in some detail in Hale, Kissock & Reiss 2007.

[15] Note that the problem is not just one of finding corresponding phonological representations in two grammars; one also must demonstrate that these are linked to corresponding semantic representations. Given normal processes of semantic shift and lexical replacement through borrowing (to use traditional terms) it is hard to imagine how the task can even be defined. For example, is Old English *mete* 'solid food' the "same word" as Modern English *meat*, which means something like 'edible animal flesh'? Common sense and standard diachronic linguistic practice say 'yes', but it is difficult to make the question coherent under the I-language approach.

primary (or, perhaps more precisely, the relevant) data for the acquirer who has constructed G_2. Note that this restricts the use of 'phonological change' to chronologically local events—there can be no more question of, e.g., whether the 'sound change' Proto-Indo-European *dw- > Armenian [jerək-] is possible (a case often discussed in the literature): it is not. The long chain of clearly distinct diachronic events which link PIE *dw to Armenian [jerək-] is not 'a change', but rather a lengthy sequence of changes. This conception of things has the advantage that it may just be possible to develop a systematic and constrained theory of sound change, thus defined, whereas it is very difficult to see what kind of coherent restrictions can be placed on sound change if PIE *dw > Arm. [jerək-] must be allowed as an example of the phenomenon.

Another significant corollary of this understanding of the nature of sound change is that the great debate in the sound change literature about whether sound changes are 'gradual' or 'abrupt' ceases to be meaningful. If change is a relationship between G_1 and G_2, and the coming into being of G_2 is dependent upon the constructor of G_2 getting his/her data from G_1, then it follows that this relationship—i.e., the 'change' event—only comes into existence the moment the acquirer's grammar G_2 does. All linguistic change is thus abrupt, indeed, instantaneous.

What does this understanding of sound change mean for 'phonologization' specifically? For one thing, there is a great deal of discussion of the process, starting in the late 19th century, which holds that the process of 'phonologization' is a 'gradual' one, transitional effects, for example, becoming steadily 'more pronounced' until such time as they get 'grammaticalized'. There are two clear senses in which this understanding of phonologization is flawed under the assumptions outlined here. First, the phonologization does not take place (by definition) until the moment the process becomes grammatically (as opposed to speech-production) driven. Whatever events precede this, if indeed there can be any, are not part of the phonologization event *per se*. Second, it is hard to see where to locate the 'gradual' developments envisioned in this scenario: we cannot place them in the grammar (else the process would already be grammaticalized), and the most restrictive theory of the *systematic* transduction systems involved in speech production (as opposed to the more 'chaotic' ones such as amount of saliva in the vocal tract) is that they themselves are constants: as unchanging systems, they can offer no account of a steady increase in coarticulation effects (as an example).

5. Synthesis: What is Phonologization?

Working with these definitions, it is apparent that phonologization cannot be the shifting of responsibility for some property of the acoustic or articulatory output from the *phonetic representation* (C) to the *phonology* (B): the representation in (C) has no properties at all except those provided by the 'phonemic level' in which underlying representations are stored and those provided by the computational system of the phonology—the phonetic representation is epiphenomenal and fleeting (not stored). 'Phonologization' must, therefore, concern itself with the shifting of responsibility for properties of the bodily output in (E) from transduction processes in (D) at one Stage, i.e., in G_1 to the phonological computation system (B) of G_2 at a later Stage.[16]

We remind the reader here that Figure 2 represents a universal model of the human 'speech chain', instantiated in each individual—this is the core of the I-language perspective. When we talk of a shift from the transduction system to the phonology, we do not mean to imply that this happens "inside" the speech chain of an individual. Rather, the difference is between tokens of speech chains that are in a diachronic relationship of immediate descent—we are referrring to Stages, as in the discussion of Hyman's work above. Thus, phonologization involves no phonological computation *per se*. Phonologization does involve a relationship between two different instantiations of level (B), two different phonologies, but this relationship is not a phonological computation.[17] Typically, as in the tonal case discussed by Hyman, the later instantiation B_2 contains a rule which was absent in the earlier instantiation B_1. Understanding phonologization also requires looking at properties of E_1, the bodily output of a grammar at Stage 1, that are explicable by reference to D_1, and seeing that acoustically similar (but "exaggerated") properties of E_2 are the result of an aspect of the phonological computation B_2 that was not present in B_1. Figure 3 sketches the model we intend.[18]

[16]Change may also occur in (A) of G_2 as a result of the analysis of G_1's (E), of course, resulting in a difference in UR between G_2 and G_1 for the 'same' lexical entry.

[17]Experience has shown that it is hard to convince people of this point. Perhaps some analogies would be useful: The relationship between two people, say, a mother and son, is not a person. The comparison between two cars ("This one is big and green, whereas that one is small and yellow") is not a car. Similarly, a relationship between two particular phonologies, yours and your mother's, is not a phonology, nor is an analyst's comparison of two particular phonologies a phonology.

[18]The properties of the human transduction systems (which we assume to be universal, thus constant) that give rise to the phonetic lowering of pitch are, of course, a property of D_2, as well as of D_1. However, the effect is potentially obscured by the fact that the phonetic representations of C_2 have voiced stops followed by representational R tones (or LH sequences) in otherwise

	Stage 1	Stage 2
Underlying Rep.	A_1: /bá/ (H tone)	A_2: /bá/ (H tone)
Phonology	B_1	B_2: Rule H $\rightarrow$ R /Voiced Stop ___
Phonetic Rep.	C_1: [bá] (H tone)	C_2: [bă] (R tone)
Speech	D_1: Phonetic lowering of pitch	D_2
Bodily Output	E_1: 🧍ba🧍 w/ pitch 	E_2: 🧍ba🧍 w/ pitch 

Figure 3: Phonologization of pitch lowering effect on vowels after voiced stops. The UR at both stages is /bá/, with a H tone. At Stage 1 pitch lowering is only a transduction effect, arising in the Speech component (D): the phonetic representation has an H tone, whose pitch is realized as slightly lowered at the beginning of the bodily output (compared with output of a UR /pá/), due to the transduction systems. At Stage 2 the (exaggerated) synchronic effect is phonological—the phonetic representation is R, due to a phonological rule. Transducing this phonetic representation leads to a pitch contour that is even more different from that of an output of a UR /pá/ because it is due to explicit 'instruction' (through featural representation) *in addition to* the unavoidable transduction effect.

Such phonologization involves the operation of the phonological learning mechanism[19] of the acquirer of G_2 on the bodily outputs produced by the speaker of G_1. The pitch lowering present in the bodily outputs is noisy, in a technical sense, and thus consistent with more than one possible source (the phonology or the transduction systems). If the acquirer misattributes responsibility for some property of the bodily output to the phonology, when, for the speaker of G_1, it was due to the transduction system, phonologization results.

At Stage 1 in Figure 3, there is a synchronic lowering of the pitch of the vowel following a voiced stop, as compared with a voiceless stop. We identify this as a lowering, only because we assume that the actual pitch differences are not encoded representationally, as an underlying tonal difference in the lexical entry of the two forms. The synchronic pitch difference between the two forms is due to properties of the transduction system—the realization of any tone in any language will be different after a voiced versus voiceless consonant. The difference is thus predictable because of assumed properties of human articulation, but you have to look outside of the grammar to see this effect—it is not represented grammatically, either in grammar inputs or grammar outputs.

At Stage 2, the grammar contains a phonological rule that changes the input tone in accordance with the phonological environment. The permanently stored tones in the two lexical items are identical at Stage 2, as they are at Stage 1, but a context sensitive rule changes the representation of the tone, say from H(igh) to R(ising) (perhaps analyszed as a sequence of Low-High) when it follows a voiced stop.

There are three "lowerings" to note when thinking about Stage 2. First, there remain the effects on pitch following a voiced stop, as compared with a voiceless one—this effect is a constant of human articulation, part of the transducers (D); this effect may be invisible if no sequences of voiced stops followed by high tones are fed to the transduction system, but it exists as a property of D nonetheless, at both stages, D_1 and D_2. Second, there is the synchronic tonal change, a lowering, from an input H to an output R, the effect of a phonological computation (B) encoded in terms of the UG-given representational alphabet for tones and the UG-given phonological computational system.

Finally, the third "lowering" is not actually a property of Stage 2, but is a relationship: the relative, diachronic lowering between Stage 1 grammar outputs

identical forms of C_1 with representational H tones. We leave this detail out of D_2 for expository clarity.

[19] Which is *not* a token of B, but rather a system not shown in our schema.

and Stage 2 grammar outputs, under the assumption that phonological lowering of H to R (due to the rule in Stage II) effects a more dramatic pitch reduction on an input H than the constant universal effects on articulation of H after a voiced stop would. This "lowering" is thus the change from pitch lowering to tone lowering from Stage 1 to Stage 2, the difference between E_1 and E_2. The first two lowerings are properties of the Stage 2 speaker. This third lowering is not a property of anything in the world other than the historical linguist's model of change, the result of an analysis.

Before we proceed, note that there is really just one difference between Stage 1 and Stage 2 speakers (with G_1 and G_2, respectively). The only difference that is not epiphenomenal is that seen in B_1 *vs.* B_2, the absence *vs.* presence of a particular operation.

6. Is Allophony Relevant?

We referred above (fn.10) to problems with the term 'phonemic' and associated notions. Let's consider a possible confusion that arises from a failure to adopt the findings of generative phonology. One result, deriving from Halle (1959), is that allophonic rules and neutralizing rules should not be distinguished—there are not two separate rule components of the grammar. In other words, a rule can be labelled allophonic or neutralizing only in the context of a particular lexicon—the difference has no status in the grammar. Now, it is tempting to suppose in discussion of phonologization that added rules are allophonic rules, and this appears to be the case in our presentation of Hyman's example from Figure 1. At stage two, the R tone appears after the voiced stop, and the H tone appears after a voiceless stop. It looks like the presence of an R is predictable from context. However, there is nothing in our sketch of the mechanism of change, the misattribution of the pitch lowering to a phonological rule that requires that the change could not take place even if (at both Stage 1 and Stage II) there are lexical R tones, including some after /b/.

The issue of phonologization is, therefore, orthogonal to the issue of whether a rule is allophonic or not. This is as we expect, since allophony, in generative phonology, has no status. It is perfectly possible that Hyman's Stage 1 has lexical items like /bá/, /pá/, /bă/ and /pă/, all of which surface distinctly at Stage I; and that Stage 2 has identical lexical items, but that the contrast between /bá/ and /bă/ is neutralized by the computational system (B).[20] The resulting pattern

[20] Of course we need to assume that there is some kind of evidence for underlying /bá/ forms, perhaps a tonal sandhi or other interaction that precedes the tone neutralization in the syn-

in Stage 2 may be allophonic, that is, predictable from context, since it may be the case that R tones appear only where an input contains a voiced stop followed by a H tone, or the Stage 2 rule may be neutralizing, in case there happen to be underlying R tones after voiced stops. The difference has to do with the content of the lexicon, and has no bearing on the mechanism of phonologization.

7. Is phonologization regular?

What would happen if the misattribution we posited as the basis of phonologization were random, rather than consistent? Perhaps phonologization applies irregularly across the lexicon. There are reasons to doubt this happens, and so we need to account for the apparent regularity of phonologization. First, sound changes are famously quite regular—at least it is easy to find lots of examples of regular ones. Second, when confronted with apparent irregularity one must be certain that diffusion, rather than direct descent, is not at issue (see Hale 2007, chapter 3 for copious discussion). Third, if the misattribution were truly random, then there would be no consistency in attribution even for tokens of individual lexical items—it would be improbable that all tokens of the word for, say, 'cat' involved a misattribution (*vis-à-vis* the Stage 1 system) whereas all tokens of the word for 'house' did not, assuming no phonological conditioning. So, truly random misattribution appears to be empirically unsupported. Fourth, recall that phonologization involves a learner building G_2 attributing to the phonological computational system of G_1 responsibility for some aspect of the bodily output which, in fact is due to the transduction system associated with the speaker of G_1. This misattribution has implications for the properties of G_2, the new grammar. Under our conception of phonology as a computation, a *function*, the mapping of input forms to output forms is regular by definition. A grammar cannot have a phonology that says "lower the tone after voiced stops in certain lexical items"—that is not phonology. If this is a good diachronic description of the stages of a "language", then we are not dealing with phonologization—some kind of lexical restructuring must have occurred to differentiate some voiced stops from others. (For present purposes, we do not need to decide if this kind of (random, non-conditioned) restructuring is possible—if it is, it is not phonologization.) If a change is restricted to a change in the phonological component—in other words, if phonologization as we have defined it does indeed exist—then it appears that the misattribution process cannot be random. Therefore, the regularity of phonologization follows from the regularity of misattribution (non-

chronic grammar of Stage 2.

regularity would require lexical restructuring—see the discussion of "phonemicization, below", so not phonologization) combined with the regularity of the computational system that is the phonology.

8. Is all sound change phonologization?

It is of some importance to recognize that phonologization does not exhaust the range of phenomena with which diachronic phonology must concern itself. Only some of the phonological learning done by the acquirer is directly based on the parse of input acoustic strings. The morphological analyses that allow learners to posit rules based on alternations require 'batch-learning' over a set of stored lexical items, which are, of course, stored using the phonological representation system provided by UG.[21] Access to such forms is not 'noisy' in the technical sense, since the forms are stored in the speaker's mind and thus are not susceptible to the same type of misparsing as raw acoustic data is. Nevertheless, the precise set of forms to which a learner is sufficiently robustly exposed within the relevant time frame certainly varies from acquirer to acquirer, and thus the batch learning will of necessity be performed on a distinct 'batch' of data for each individual, even those exposed to the output of the same source. Since the data will be different, it follows that the generalizations over the data may also differ—to make this concrete, each son of a mother who provides PLD will end up with a distinct I-language. A difference in the productive morphological parses assigned to the data may, in turn, entail a difference in the set of phonological rules which the acquirer must posit in order to synchronically relate allomorphs. This would be phonological change, but it would not be phonologization, as defined in this chapter, since the mechanism of change need not involve the elevation of phenomena originally triggered by the transduction systems to the phonological level.

The process of 'dephonologization', to which we turn in the next section, is also, strictly speaking, not 'phonologization', obviously. It does, however, bear a close relationship to phonologization, to the point that it seems worthwhile to discuss it briefly here.

[21] If the first word a learner hears is *cats* [kʰæts], there is no way of knowing whether the initial velar stop is a plural marker, or if the aspiration on that stop is a marker of animacy. An analysis depends on comparison of multiple forms. When the learning extends to additional forms like *cow* [kaw] and *cows* [kʰawz], the learner will be able to identify a distinction between root and plural marker, relating the [z] and [s] plurals as a single underlying form and constructing a phonological rule that generates the alternants from this one form. All this learning requires an initial stage of storage of unanalyzed forms.

9. Dephonologization

In Proto-Indo-European an allophone of *t arose before a following syllable-initial *t, of the form *t^s.[22] This presumably resulted from the phonologization of the transition between the two dentals across the syllable juncture[23]—the addition of a rule to the phonological component of the grammar. This rule (or its reflex, since in the historical period, [t^s] becomes [s]) survives into many of the Indo-European daughter languages, including the members of the Iranian branch – the closest relative of the Indic branch to which Sanskrit belongs. However, Sanskrit does not have the rule, i.e., the outcome of late common Indo-European *Vt^{s}tV is *VttV. The allophony-triggering rule was lost from the grammar, and with it the allophony. In the history of the Indic branch, the prior phonologization has been lost. There has been dephonologization. Presumably, the stop transitions of *tt* clusters in the pre-phonologization stage and the Sanskrit post-phonologization stage are identical, since they are due to the universal transduction systems. At some point, in the Sanskrit linguistic tradition, learners misattributed the high frequency noise between the stops to the transducers (D in the chain), instead of to the presence of a feature for stridency in the phonetic representation (C). So, dephonologization is also a kind of sound change.[24]

10. Rethinking Phonemicization as Lexification

Consider Hyman's example in Figure 1 again repeated as Figure 4 below. How do we understand the shift from Stage 2 to Stage 3? Looking at the B step in the chain, it seems obvious that B_3 will not have the rule lowering H to R that is present in B_2—the relevant tones are stored as R (in A_3), so there is no need for a rule to make H into R.

However, there are not one but *two* differences between A_2 and A_3—the tones are different and the consonants are different. This double difference is necessary in the context of a discussion of 'phonemic' changes. It is only because

[22] Similar developments affected the other dentals in heterosyllabic dental+dental clusters; we consider *t alone simply for expository ease.

[23] Whatever its precise cause, it must have arisen via the reanalysis of the perceived acoustic properties of this juncture, there being no morphological analysis type of explanation available. To be honest, we expect no transition here, just the long closure typically associated with geminates, but our main point is that changes appear to be reversible—misattribution can happen in both directions. Perhaps the development to *t^{s}t* must be taken as evidence that the first stop in the cluster was, in fact released.

[24] See Dresher (this volume) for an example of this type from the history of English as well as for discussion of some of the same issues raised in this paper.

	Stage 2	Stage 3
Underlying Rep.	A_2: /bá/ (H tone)	A_3: /pă/ (R tone)
Phonology	B_2: Rule H $\rightarrow$ R /Voiced Stop ___	B_3
Phonetic Rep.	C_2: [bă] (R tone)	C_3: [pă] (R tone)
Speech	D_2	D_3
Bodily Output	E_2: 🧍ba🧍 w/ pitch ⌣	E_3: 🧍pa🧍 w/ pitch ⌣

Figure 4: "Phonemicization". This should be broken down in two stages: (1) lexification and loss of the rule; (2) merger of /b/ and/p/.

the initial consonant in A_3 is now non-distinct from, for example, another lexical item, [pá], that the R vs. H distinction has become phonemic, unpredictable. Since we have already decided to be skeptical of the notion 'phonemic', let's delve deeper.

There is no reason that the H-R contrast could not have been "phonemic", that is lexically contrastive, even at Stage 2. Hypothetically, Stage 2 could have had lexical items /bá/, /pá/, /bă/ and /pă/ with a neutralizing rule that mapped inputs /bá/ and /bă/ both to [bă] (see fn. 20). Just the loss of this rule, with its effects transferred into the lexicon, would yield a Stage 2b with lexical entries /pá/, /bă/ (occurring with two meanings) and /pă/.[25] The voicing contrast remains, as does the (hypothesized) tonal contrast (after /p/), so there is nothing to call "phonemicization". Whether this exact change is plausible or not, any case of partial merger provides the same logical structure. This change is sketched in Figure 5. Since phonemicness is not relevant to the phonology (see discussion of allophonic versus neutralization rules above), we don't need to indicate the full range of possible lexical items in Figure 5. The change, in I-language terms, given our model of the speech chain, is not affected by whether or not the languages happen to have lexical forms like /pă/ or not.

We conclude from such examples that phonemicization as illustrated by Hyman's example, is actually a combination of two processes that need to be unpacked in order to understand what is going on. The loss of the phonological

[25]For example, whatever provided evidence for an underlying /bá/ might disappear, forcing an analysis of all surface [bă's] as originating exclusively from /bă/.

	Stage 2	Stage 2b
Underlying Rep.	A_2: /bá/ (H tone)	A_{2b}: /bă/ (R tone)
Phonology	B_2: Rule H $\rightarrow$ R /Voiced Stop ___	B_{2b}
Phonetic Rep.	C_2: [bă] (R tone)	C_{2b}: [bă] (R tone)
Speech	D_2	D_{2b}
Bodily Output	E_2: 🚹ba🚹 w/ pitch ⌣	E_{2b}: 🚹ba🚹 w/ pitch ⌣

Figure 5: "Lexification". The R tone is derived from H at Stage 2, but stored in the lexicon at Stage 3. The Stage 2 tone rule is not part of the Stage 3 phonology.

rule and representation of R tones for Stage 2 H tones we will call *lexification*. This parallels our discussion of phonologization in that responsibility for an aspect of speech generation shifts from one component to another between Stages, here from the phonology to the lexicon. The possibility of a subsequent merger of /p/ and /b/ to yield Hyman's Stage 3 is not a logical necessity—it is clear that language change does not have to maintain the number of underlying contrasts from Stage to Stage, although this is perhaps implied by Hyman's example. If such a merger did occur, it would be *just* a difference between the lexical representations of the stages—it might involve a shift in "phonemic" contrasts, but since these have no status in the modern model of grammar, they deserve no status in a model of the relationships between grammars in the direct descent relation, *aka* language change.

In old fashioned terms, the lexification of R tones only becomes phonemic through the merger of p and b, but this merger is not an inevitable consequence of the occurrence of the change sketched in Figure 5. Stage 2b is a perfectly reasonable language, with or without lexical /pă/.

To get from Stage 2b to Stage 3 we need something like the change sketched in Figure 6, a merger of /p/ and /b/. As mentioned above, such a change might have the effect of making the H versus R distinction phonemic, or it might not—that depends on whether the lexicon already had elements like /pă/.[26]

[26]That is, at Stage 2b, the lexicon had elements like /pă/ that were accurately transmitted to Stage 3.

	Stage 2b	Stage 3
Underlying Rep.	A_{2b}: /bă/ (R tone)	A_3: /pă/ (R tone)
Phonology	B_{2b}	B_3
Phonetic Rep.	C_{2b}: [bă] (R tone)	C_3: [pă] (R tone)
Speech	D_{2b}:see text	D_3
Bodily Output	E_{2b}: 🕴ba🕴 w/ pitch ⌣	E_3: 🕴pa🕴 w/ pitch ⌣

Figure 6: Merger of /b/ and/p/. Stage 2 /b/ corresponds to Stage 3 /p/. The existence of contrast is irrelevant to a characterization of the nature of the changes—Stage 2 may or may not have /p/, as well as /b/.

11. Relationship of Phonologization and Lexification

We have already argued that lexification need not be followed by a merger of the type in Figure 6: a language tradition can change from Stage 2 to Stage 2b and "stay there". There is no mechanism for looking ahead at Stage 2 and only changing to Stage 2b if it is guaranteed that Stage 3 will follow, since each Stage is instantiated in different minds.

We offered a reinterpretation in Fig. 5 of the term phonemicization (now *lexification*) as a relationship between a rule-generated variant at an earlier Stage corresponding to a stored form at the subsequent Stage, a shift of responsibility from the computational system (B) to the lexicon (A). Let's refer to this change as "Indirect lexification"for reasons that will immediately become apparent. This contrasted with our discussion of phonologization which involved a shift of responsibility from the Speech step (D) to the Phonology (B), shown in Fig. 3.

There is another possibility that we relied on implicitly in Fig. 6. Note that this corresponds to an unconditioned change from /b/ at Stage 2b to /p/ at Stage 3. It appears to be the case, although we are arguing from a *lack* of evidence, so we cannot be certain, that such unconditioned changes can, in fact occur. In other words, lexical representations appear to be able to shift between Stages without an intermediate Stage involving a phonological rule. If there is no such stage involved, then it must be the case that the bodily output (E) at the earlier stage can be misanalyzed directly as a difference in lexical representation. Let's call this "Direct lexification".

We can now express the changes presented in Hyman's example as a sequence

of phonologization (call it Output-to-Phonology), followed by Indirect Lexification, followed by Direct Lexification. In this particular case, the effect of the Direct Lexification was to merge a contrast (/b/ versus /p/) that had conditioned the Indirect Lexification. However the mechanism of change is not logically related to its effects on the system of contrasts present underlyingly or on the surface. This means that Direct Lexification should be able to occur whether or not it results in a merger, and whether or not it is preceded by a "related" Indirect Lexification.

The unconditioned change of the Indo-European voiceless stops to fricatives in Germanic is an example of an apparent Direct Lexification change, involving no merger.[27] In contrast, the unconditioned change of Indo-European voiced aspirates to plain voiced stops *did* result in a merger with inherited voiced stops in the Balto-Slavic family. So, Direct Lexification changes may or may not result in merger themselves.

How does Direct Lexification occur? We assume that it is quite possible that an initial sample of hits for a given segment with a skewed distribution (predicted, occasionally, if the distribution is stochastic[28]) may lead to a reassignment of the target space for that segment, and thus a direct Direct Lexification "rephonemicization"—that is G_2 encodes segments in the new lexicon with a different set of features from the "corresponding" segments in the G_1 lexicon—without a phonological rule ever having been in play. In fact, it is hard to motivate the kind of unconditioned phonological rules that such a scenario would require—why would a learner posit a lexicon full of voiced aspirates along with a rule that changes them *all* into plain voiced stops? Either the learner would perceive the voiced aspirates in the Primary Linguistic Data, store them and produce them, or s/he would not perceive them as such, markedness and repair based models of change notwithstanding (see Scheer, this volume).

There is also no reason to suppose that a Direct Lexification needs to be preceded by phonologization (Output-to-Phonology). The unconditioned development of Indo-European *a to Sanskrit /ə/, shows no evidence of a stage involving 'allophony' or the 'exaggeration' of locally conditioned coarticulation effects which are taken to characterize phonologization events.[29]

[27]There may have been a stage with aspiration, but the focus here is that there was no merger and there is no evidence for a rule in any synchronic grammar.

[28]This stochasticness reflects the interaction of all the systems involved in the Speech level—there is no reason to posit a stochastic aspect to the grammar itself.

[29]Indo-European *e and *o also changed to Sanskrit /ə/, without any direct compensation for the loss of contrasts seen in Hyman's illustration—sound changes can't look ahead to evaluate

	Stage 1	Stage 2′
Underlying Rep.	A_1: /bá/ (H tone)	$A_{2'}$: /pá/ (H tone)
Phonology	B_1	$B_{2'}$
Phonetic Rep.	C_1: [bá] (H tone)	$C_{2'}$: [pá] (R tone)
Speech	D_1: Phonetic lowering of pitch	$D_{2'}$: "Invisible" lowering of pitch
Bodily Output	E_1: 🕴ba🕴 w/ pitch ◜‾	$E_{2'}$: 🕴pa🕴 w/ pitch ‾

Figure 7: Merger without lexification. There is no relevant difference at step B between the Stages. The only difference of any import is at the lexical level A. The differences at C and E are all derivative of the lexical difference. Of course, pitch lowering after voiced stops remains a constant property of the speech level (D), but it is irrelevant once those stops are devoiced. This situation involves a difference in lexical representations, but not the lexification of the effects of a rule.

We have seen that Hyman's Stage 1 can develop to Stage 3, but it can also remain at Stage 2 indefinitely—Stage 2 is a possible language, showing just phonologization from Stage 1. Stage 2b is a possible language as well, showing phonologization, followed by lexification. Stage 3, in which merger of the stops follows Stage 2b is also possible, representing what Hyman calls phonemicization. However, the merger that is necessary to phonemicization could occur without lexification. For completeness we sketch this possibility in Figure 7. This situation involves a difference in lexical representations, but not the lexification of the effects of a synchronic Stage 1 rule. The development from Stage 1 to Stage 2′ in Figure 7 cannot lead to Hyman's Stage 3 since the distinction between Stage 1 /pá/ and /bá/ has been irrevocably lost. We thus see that lexification and merger which combine in Hyman's phonemicization scenario are logically distinct and completely independent of each other: either can occur without the other.

In addition to illustrating the possibility of dephonologization, the Sanskrit *tt* case bears on the putative relationship of phonologization and phonemicization, and the notion of drift—the idea that sound changes in a "language" tend

their results.

in certain directions. Why do we even raise this issue? The relationship between phonologization and phonemicization is sometimes discussed under the general heading of the 'life cycle of phonological rules', the idea being that such rules are 'born' via the mechanism of phonologization, live for some span as phonological rules, and then 'die' via the mechanism of phonemicization. Indeed, Hyman (1976) explicitly ties the two phenomena together, noting that "*the development* [via phonologization—mh,mk&cr] *of a phonological rule carries the seeds of its own destruction*" (emphasis in original). While engagingly phrased, this represents, we fear, a rather coarse oversimplification of the matter. While birth provides one of the necessary antecedent conditions (life) for death, phonologization is not necessary for the coming into being of a phonological rule (since these may arise via morphological parsing, as discussed above), and thus it is entirely possible that many cases of "phonemicization" involve the loss of phonological rules which have simply nothing to do with earlier phonologization events. And while birth, unfortunately, entails death as a necessary consequence, phonologization in no way entails phonemicization as a necessary consequence. This is clear from the Sansrkit example discussed above, where the phonologization was just reversed—something like a Stage I to Stage 2 shift does not have to be followed by something like a Stage 2 to Stage 3 shift. In the Sanskrit case, we seem to revert back to Stage 1 from Stage 2. Finally, we reiterate that a merger of a distinction can occur between Stages without a concomitant lexicalization of another contrast—losing the voicing contrast, but gaining the tonal contrast. We suspect that the issue only arises due to an abiding crypto-functionalism in the field that expects, in the face of plentiful evidence, that functionalist considerations like "functional load" play some kind of role in either synchronic or diachronic linguistic explanation. (See Hale & Reiss (2000, 2008) for our views on functionalist reasoning.)

Hyman's discussion of "life cycles" is part of a long-standing discussion of "drift" in language change. The notion is discussed by Jespersen, Jakobson, Sapir and more recent scholars, often in a manner that obscures the relationship to earlier work, for example, with claims that language change is a process of optimization or markedness reduction. However, there is no mechanism in the I-language view for the propagation of drift from one generation to another. A learner has no access to previous Stages of his/her linguistic tradition. The learner can only access the outputs of individuals in his/her environment. If there is a difference between the learner's phonology/lexicon and the phonologies/lexicons of those who provided the PLD outputs (a 'change'), then the learner must have misanalyzed that output. Note that discussions of drift and life cycles

never make precise claims about the time scales at which these forces work, and this must be for the simple reason that the claims are based on entities, "languages" like English, Sanskrit or Cree, that are non-existent under the I-language approach. How would the mechanism of drift be encoded in the grammatical knowledge of a speaker? Why did Sanskrit "drift" back to an earlier Stage while Iranian did not?

We find that the inherited rule affecting Indo-European *tt* clusters was synchronically productive in Proto-Iranian and (with a slightly modified output) in the Old Iranian languages, since its most frequent context of application was across morpheme boundaries, when *t-initial affixes (of which there were many) were added to dental-final roots (ditto). Since those dental stop-final roots also normally appeared before non-dental-inital suffixes, in which context their dental stop was preserved, the rule remained synchronically motivated. This rule, which, arose via phonologization, has persisted for these 3,000 years without triggering any phonemicization event into the modern Iranian languages. Maybe someday it will, and the 'life cycle' will find its expression in Iranian (but for how long can one say that a diachronic process is 'in transition'?): but, in any event, it never will, and can't, in Sanskrit (or its descendants).

It seems clear, then, that the two events, phonologization and lexification (the old 'phonemicization'), are simply independent phenomena of diachronic development with merger falling out epiphenomenally depending upon idiosyncratic properties of lexica. As independent phenomena, they may appear to interact, but this "interaction" has no ontological status—it is an artifact of our point of view as analysts. The 'life cycle' of phonological rules is a mirage—implying causal linkage where none exists. Understanding this allows us to focus our research attention on clarifying the actual factors which give rise to each of these processes, and to turn away from a search for the mystical, pan-generational forces which would need to be in play to enforce a 'life cycle'.

12. Description vs. Explanation of Change

Finally, we would like to clarify the distinction between the *description* of language change and the *explanation* of language change. Our model of the speech chain included (A) underlying representations (we didn't treat the thorny issue of how lexical entries are combined into complex forms) which are the input to (B), phonological computation, yielding (C) the output phonetic representation which is fed to a complex system of transducers (D), finally yielding the acoustic output of the body (E). The contents of the lexicon that appears in A is

subject to variation, but all content is encoded in a single universal alphabet of features. The contents of the computational system (B) is also subject to variation—languages appear to have different computations/processes. None of the other components (C,D,E) are subject to variation: we assume that the transduction systems (D) are constant, and the phonetic output (C) and bodily output (E) are just the fleeting products of the higher steps in the chain. For example, if two languages G_1 and G_2 differ with respect to the respective forms c_1 and c_2 at (C) in the speech chain, then the difference must be due to either a difference in the lexical input for c_1 versus c_2, or a difference in the computation, or both. Different inputs to the phonology can give different outputs; identical inputs (A-level) to the phonology can give different outputs only if the computational system B_1 and B_2 differ appropriately.

If the speakers of two I-languages have two different forms at E, e_1 and e_2, respectively, then the difference is due to one of two sources (or both). One case is that in which there is a difference in what was received by the invariant transduction system, in other words, a difference at step (C). As we have seen, this regresses back to a difference at either A or B or both. The other case is where a difference is introduced in D. This is possible, despite the fact that D is assumed to be universal, because D has other inputs than just C, a fact alluded to above, but somewhat obscured by Figure 2. To give a simple example, the decision to scream a word, rather than whisper it, changes the (non-grammatical) input to the transduction, but does not constitute variation in the transduction system. Intention to scream is not represented at C, but it does affect D by some other (non-linguistic) path.

So, the *description* of a language change (a lack of correspondence between grammars) can only make reference to differences in components (A) and (B)—these are the only components that are subject to change and they are the only components that are linguistic. (C may be considered linguistic, but it is derivative from A and B. E is not linguistic—it includes the effects of shouting or talking with a cigarette dangling from the lips, and it too is derivative.)

However, the *explanation* of a language change must make reference to more than grammar. The batch learning morphological parsing of initially unanalyzed stored forms mentioned briefly above is not part of grammar, but rather part of the learning process that *builds* a grammar by inducing lexical entries and phonological rules (the content of A and B). Similarly, the parse of bodily output by the G_2 acquirer, in which acoustic properties have to be attributed (rightly or wrongly in light of the noise) to C_1 (the phonetic representation output by G_1) is itself not grammar, but transduction, where the acquirer must assign a featural

representation to a signal originating in the auditory system.

13. Conclusion

We hope to have shown that (1) carefully maintaining the I-language perspective; and (2) establishing explicit definitions for change, grammar, transduction and the like, brings clarity to a family of phenomena studied by historical linguists. Most fundamentally, drawing on the fuller treatment of Hale 2007, we suggest that language "change" is something of a misnomer, since we assume that languages, that is I-languages, do not change (other than additions to the lexicon) after maturity. We pointed out that language change is by definition punctual under our approach. If G_2 is constructed on the basis of output of G_1, then the change comes into being by virtue of the construction of G_2, say, as soon as a given rule is constructed as part of G_2. We embedded our model of language, that is grammar, inside a model of the speech chain that included the mapping from lexical forms in long-term storage to the sound waves that constitute bodily output. What is traditionally called language change, is under our view, a shift—between stages of a linguistic tradition (a sequence of I-languages in which each is the direct descendant of its immediate predecessor with respect to the traits in question)—in which component of the speech chain (of Figure 2) is primarily responsible for aspects of the bodily output (E of the chain).

References

Chomsky, Noam. 1986. *Knowledge of Language*. New York: Praeger.

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

Dresher, B. Elan. Rule-based generative historical phonology. This volume.

Hale, Mark. 2007. *Historical Linguistics: Theory and Method*, Oxford: Wiley-Blackwell.

Hale, Mark. 2003. Neogrammarian Sound Change. In Joseph & Janda 2003.

Hale, Mark, Madelyn Kissock & Charles Reiss. 2007. Microvariation, Variation, and the Features of Universal Grammar. *Lingua* 117.4 383-415.

Hale, Mark & Charles Reiss. 2008. *The Phonological Enterprise*. Oxford University Press.

Hale, Mark & Charles Reiss. 2000.Substance abuse and dysfunctionalism: Current trends in phonology . *Linguistic Inquiry* 31: 157-169.

Halle, Morris. 1959. *The sound pattern of Russian*. The Hague: Mouton.

Hammarberg, R., 1976. The Metaphysics of Coarticulation. Journal of Phonetics 4, 353-363.

Hyman, Larry. 1976. Phonologization. In *Linguistic Studies offered to Joseph Greenberg*, A. Juilland (ed.), Anma Libri: Saratoga, CA, 2:407-418.

Hyman, Larry. 2008. Enlarging the Scope of Phonologization. *UC Berkeley Phonology Lab Annual Report* pp. 382-407, Berkeley: UC Linguistics.

Joseph, Brian & Richard Janda. 2003. *The Handbook of Historical Linguistics*. Oxford: Blackwell.

Keating, Patricia. 1988. Underspecification in Phonetics. *Phonology* 5: 275-292.

Kingston, John. Tonogenesis. In M. van Ostendorp, C. Ewan, E. Hume, and K Rice (eds.) *Companion to Phonology*. Oxford: Blackwell Publishers. (To appear)

Kingston, John. 2007. The Phonetics-Phonology Interface. In Paul de Lacy (ed.), *The Cambridge Handbook of Phonology*, pp. 401-434.

Kingston, John and Randy Diehl. 1994. Phonetic Knowledge. *Language* 70: 419-454.

Kiparsky, Paul. Phonologization. (this volume)

Matthews, P. H. 1998. *The Oxford Concise Dictionary of Linguistics*. Oxford: Oxford University Press.

Scheer, Tobias. Crazy rules, regularity and naturalness in diachronic and synchronic phonology. This volume.

Tatham, M. A. A. (1971) Classifying allophones. Language and Speech 14, 140-145.