On Explaining Opaque Sound Change: Potential Counterexamples to Phonetically Motivated Change and their Consequences

David Constine

1 Introduction

Throughout the history of historical linguistics as an academic discipline, sound change has been explained in various ways and to varying degrees of believability. While nineteenth century explanations for sound change were at times laughable (for instance, the idea that languages spoken in northern climates were predisposed to certain types of change due to a preference to keep one’s mouth closed as to conserve heat, Falconer 1781), certain ideas have withstood the centuries, and continue to inform modern explanations of sound change. Chiefly, this is the idea that sound change has some sort of phonetic basis - be it articulatory or acoustic, though nineteenth century scholarship typically favored articulation. Today, most academic frameworks explaining sound change appeal to the phonetic properties of speech. While phonetic explanations are often elegant, and seem strong enough to account for many instances of regular sound change, it is not always clear that they can account for all instances of regular sound change. Phonetically opaque (or “unnatural”) sound changes are certainly not unheard of - even to the linguists of the nineteenth century, for example, the relatively opaque ruki rule of the satem Indo-European languages. Given the unprecedented breadth of language data available today, it is unsurprising that opaque changes continue to be identified. Several curious examples of opaque change are presented by Robert Blust in his 2017 paper “Odd Conditions: Context Sensitive Sound Change in Unexpected Places”. Coming from a wide array of Austronesian languages, Blust’s data presents a curious puzzle for those looking to maintain that sound change
is phonetic. Over the course of this thesis, I will put a portion of Blust’s data in conversation with several theories of phonetically motivated change, and ultimately show that while Blust’s data is certainly difficult to account for phonetically, phonetic explanations are not incompatible with his data and seem to be generally more supportable than Blust’s alternative: conscious, deliberately induced language change on the part of the speech community.

1.1 Definitions and Context

It is important to begin by describing precisely what sorts of sound changes are relevant to this discussion. The sorts of sound changes to be considered are those that seem to have the same character of phonetically motivated internal changes in the sense of their propagation, but lack any apparent phonetic motivation (at least, according to Blust). This is to say that the changes to be considered are suggested to have no clear articulatory, acoustic, aerodynamic, or perceptual explanation for their occurrence. At the same time, the relevant sound changes lack any apparent evidence to support an explanation by language contact or internal analogy. Additionally, these changes are non-sporadic, but rather are regular, affecting the language systematically in all instances in which the relevant change could occur.

1.2 Theories of Phonetically Motivated Change

This section will outline three different, though not totally dissimilar, theories of phonetically motivated change: John Ohala’s model of Hypocorrection and Hypercorrection, Juliette Blevin’s model of Evolutionary Phonology, and Andrew Garrett and Keith Johnson’s model of Bias Based Change.
1.21 Hypocorrection and Hypercorrection

In Ohala’s paradigm of sound change, the two primary motivators for change are hypocorrection and hypercorrection, two processes rooted in speech production and perception respectively (Ohala 1993). Crucially, both revolve around “correction”, a proposed grammatical mechanism through which listeners address and “repair” instances of interference in the phonetic signal in order to maintain shared internal representations of a language within a speech community. Hypocorrection and hypercorrection, and consequently sound change itself, are manifestations of this mechanism either failing to be applied, or being applied incorrectly.

Hypocorrection is a phenomenon during which the speaker makes a speech “error” due to physiological factors involved in speech production, which then goes on to be internalized by listeners, resulting in a regular sound change. This is most frequently understood in the context of assimilation. Ohala provides the common cross-linguistic example of a VN segment becoming a nasalized vowel (/VN/ → /N/). In Ohala’s understanding, due to the physiology of speech, vowels preceding nasals are always nasalized to some degree (though may not always be perceived this way). Consequently, if overt articulation of the nasal fails to occur, one is left only with a nasalized vowel. While the listener may typically “correct” these utterances (for example, even though the nasal might not at times be fully realized, the listener still perceives the speech signal to include the nasal, even if the signal itself closely resembles a nasalized vowel in its production and acoustics), in an instance of hypocorrection, the intended form of the speaker and the understood form of the listener no longer match, thus yielding a nasalized vowel with no following nasal consonant.
Hypercorrection, on the other hand, is born entirely from the listener in a process that is essentially opposite to hypocorrection. Hypercorrection relies on the listener to erroneously "correct" a speech signal in a manner that is not consistent with the speaker's intention. In this case, the listener misinterprets a segment to be the result of some type of assimilation, insertion, or deletion, rather than a realization of the speaker's intention. The listener then internally dissimilates this segment in an attempt to "correct" the influence of articulatory interference. This interference, however, never actually took place and the ensuing dissimilation becomes a novel sound change.

An example of hypercorrection given by Ohala is the dissimilation of front vowels in palatal environments. Phones adjacent to palatals often receive some sort of phonetic interference. In the case of vowels, this involves some sort of fronting. While a non-fronted vowel in a palatal environment might have acoustic similarities to a front vowel, a listener will typically be able to tease out the intended phone through normal correction. However, if the listener applies this same type of heuristic when a front vowel truly was intended by the speaker, this results in a dissimilation via hypercorrection.

In both hypo and hyper correction, the listener is essentially either failing to correct, or erroneously correcting for the effect of coarticulation coming from the speaker. These phenomena would both seem to be tied to categorical perception, the discreet perceptual distinction between two related phones or segments that differ only in one crucial variable that exists on a continuum.

To Ohala, through hypo and hyper correction, both natural and unnatural changes are the results of perceptual phenomena rooted in phonetics:
"The account given here provides a consistent, integrated account of sound changes generally regarded as opposite in character, "natural" vs "unnatural" assimilative vs. dissimilative. Central to this is the distinction between hypocorrection where the listener copies at face value the naturally occurring perturbations in the speech signal, thus producing natural, assimilative change, vs hypercorrection, where the listener unnecessarily corrects the speech signal, thus giving rise to unnatural, dissimilative changes" (Ohala 1993, 265).

Yet, can Ohala’s framework account for all instances of unnatural change? It should be noted that Ohala’s conception of "unnatural" seems rather limited. While dissimilation may seem unnatural on the surface, following Ohala it would seem fairly simple to trace its origins back to phonetics. Not every instance of unnatural change is so transparent, nor so easily categorizable, as will be seen in Section 2.

1.22 Evolutionary Phonology

Evolutionary Phonology (Blevins, 2004) is a theory of phonology which understands typologically common synchronic sound patterns as the result of "recurrent" sound change - sound change that is typologically common due to phonetic factors (speech production, perception, and acoustics). Within its account for the origins of sound change, Evolutionary Phonology presents three sound change classifications referred to as CHANGE, CHANCE, and CHoice. CHANGE involves listener misperception of a segment due to similarities in the phonetic properties of a spoken utterance with those of the utterance as perceived by the listener. For example, while a speaker might say [anpa], this utterance could be misheard and interpreted as [ampa] by the listener due to perceptual phenomena involving nasal-stop clusters, in which place cues are much stronger in the stop than in the nasal (Blevins 2004). As such, sound changes due to CHANGE rely on an error in speech perception.
CHANCE differs from CHANGE in that rather than amounting to a perceptual error, sound changes in this category result from ambiguities inherent to the signal itself. Dissimilations (among other phenomena) are explained through CHANCE. Thinking back on the dissimilation of front vowels in palatal environments, one can understand this segment to have multiple interpretations not due to misperception per se, but to perceptual ambiguity. As such, the segment [ej] could reasonably represent an intended utterance of /ej/ or of /əj/, depending on precisely how one interprets the fronting caused by the following palatal. If listener perception, for whatever reason, fails to match the ambiguous signal to the speaker’s intention, a sound change occurs due to CHANCE.

Changes due to CHOICE occur due to variation in pronunciation of a relevant segment. If a listener is exposed to multiple, perceptually distinguishable forms of the same utterance, this will lead to the creation of an exemplar form within the listener. If this form differs from that of the speaker, then this will result in a sound change by CHOICE. Here, Blevins gives an example from Bellonese. The Bellonese word /ka-kata/ to laugh has several variants depending on register - [kakata] in hyperarticulated speech, [kkata] in fast speech, and an intermediate form [kákata]. If one hears [kkata] most frequently (thus being the form with the greatest token frequency), this (as well as identical segments in other words) will likely become the internalized form, leading to a sound change by CHOICE.

1.23 Bias Based Change

The account of sound change presented in Garrett and Johnson (2012), which I refer to as Bias Based Change, is a deductive theory of sound change drawn from observable biases in speech production and perception. Rather than constructing a theory from instances of sound change
directly, Garrett and Johnson seek to establish how linguistic biases may lend themselves to certain varieties of change.

Bias Based Change is built around the idea that inherent in language usage are bias factors. Bias factors arise due to the fact that certain variations in speech production and perception are more likely to occur than others, (that is to say, the occurrence of articulatory and perceptual variations is asymmetric), and that this variation is due to four overarching phonetic processes: motor planning, aerodynamic constraints, gestural mechanics, and perceptual parsing.

Garrett and Johnson define motor planning as the “process of constructing or retrieving motor plans that will later be executed by speaking” (Garrett and Johnson 2012, 15); essentially, motor planning is the cognitive process that precedes and structures articulation. The two primary errors in motor planning that Garrett and Johnson appeal to are blending and inhibition.

Blending refers to situations in which nearby, similar segments influence each other over the course of activation. This can result in various sorts of assimilation, as well as spoonerisms - for example, errors such as snow flurries > flow snurries, or reading list > leading list. Crucially, motor plan blending acts on segments that are “adjacent” to each other in some way. This could refer to literal adjacency in an utterance, but also to adjacent syllable/word position (i.e, interactions between onset or nucleus segments with other onsets/nuclei), or to similarities in phonetic structure between segments. Inhibition, on the other hand, refers to the tendency to prefer an alternating pattern in onset structure (an ab pattern), rather than one which repeats onsets in an aabb pattern. This sort of motor planning error is frequently exemplified in tongue twisters. Repetition of the phrase unique New York will typically result in utterances of [junik ju nørk] or [nujik nu jørk].
Sound change caused by motor planning bias factors are understood as speech errors that have been phonologized into the language. While Garrett and Johnson maintain that sound changes that have their origin in speech errors are rare, they are not impossible. Specifically, instances of non-local assimilation and dissimilation seem to be rooted in motor planning error bias factors. For example, the process of consonant harmony in Navajo strongly resembles instances of blending:

(1) Navajo (Athabaskan) sibilant harmony (McDonough 1991, cited by Hansson 2010, 44)
   a. /ʃ-\-iʃ-mas/ > [jimas] ‘I’m rolling along’
   b. /ʃ-\-is-ná/ > [sisná] ‘he carried me’

Instances of non-local metathesis (i.e., of the variety exemplified by the change of Latin castrum “fort” to Old Sardinian crastu) also show a blending pattern.

Aerodynamic bias factors are those that “are characterized by a tendency toward phonetic change as a result of changing aerodynamic parameters even when all else (e.g. the position of the articulators) remains constant” (Garrett and Johnson 2012, 17). This bias exists due to two crucial laws of the aerodynamics of speech. The first is the aerodynamic voicing constraint - the fact that vocal fold vibration depends on a greater air pressure below the glottis than above it. Following from this arises a hierarchy of ease for voicing certain manners of articulation, with stops being the most difficult to voice. The second law of speech aerodynamics that Garrett and Johnson appeal to is the fact that frication requires sufficient air pressure behind the place of articulation. As such, voiced fricatives (all else being equal) are more likely to become glides due to the reduced air pressure brought on by voicing.
Bias factors arising from gestural mechanics are those that come from the movements of articulators, and the interactions between them. Garrett and Johnson provide two classifications for these bias factors, gestural overlap and gestural blend. Gestural overlap involves at least two separate articulators moving simultaneously. If the motion of one articulator obscures that of another, this will result in gestural overlap. One example of gestural overlap in speech can be seen in the pronunciation of the phrase “hand grenade”. Frequently, this phrase is realized more closely to [hæŋ ɡəneid] than as [hænd ɡəneid] due to a timing mistake that articulates [g] before [d] (Cho 2011, cited in Garrett and Johnson 2012, 19). This results in variants of the phrase “hand grenade” that seem to lack the coronal sequence [nd]. Should this phenomenon persist across similar utterance, this might then be internalized by the listener and result in a sound change.

Gestural blend, in contrast, results from competing demands on a single articulator. An example of this is the fronting of the velar stop [k] when preceding a front vowel. Here, the tongue body is required to move backward and forward in the mouth in a very short amount of time. The fact that [k] becomes fronted in this environment is the result of the blending between the tongue’s two required places of articulation.

Finally, biases in perceptual parsing are those that are grounded in perceptual misunderstandings of the phonetic signal. Garrett and Johnson maintain that biases rooted in perceptual parsing are the least understood of the bias factors involved in their theory of sound change. While they allude to general tendencies for lax vowels to lower when they are misperceived, or that certain acoustic features are more likely to be misperceived than others, they hold that more research is needed before these biases can be adequately explained.
One interesting example of sound change due to a perceptual parsing bias also involves an instance of gestural blend. The change of \( *k > \empty v / _V[^{\text{high}}] \) in Old English can be explained first by gestural blend causing the fronting of [k], and its subsequent affrication due to a perceptual bias that causes [ki] segments to be frequently misheard as [\empty v j]. This misinterpretation was eventually phonologized, thus bringing this specific sound change.

2. Phonetically Unmotivated Changes in Austronesian

While the three theories of phonetically motivated change presented by Ohala, Blevins, and Garrett & Johnson all provide powerful explanations for why a great number of attested internal, regular sound changes happen, what does it mean for these theories if there is data that is incompatible with all of them? It is precisely this sort of data that is presented in Blust (2017). Blust provides twelve examples of historical changes in languages of the Austronesian family that all seem to lack any apparent phonetic motivation. Crucially, they do not seem to be the result of language contact, and for the most part, they happen systematically throughout the languages in which they occur. The only difference, then, between Blust's sound changes and those modeled in the three theories previously discussed, is their phonetic opacity.

It is important to note, however, that Blust's data is not entirely unproblematic. Of the twelve changes he discusses, several are mappings of proto-forms onto reflexes for which the distance between them spans, at times, millennia. Given enough time, the relationship between any reflex and its proto-form could end up appearing quite opaque; this does not necessarily mean that the path leading to this situation lacks a phonetic explanation. Still, with this being said, the path from Blust's proto-forms to his reflexes is unattested, and while a phonetic path can be proposed, it most likely could never be supported with direct evidence. Consequently, any
analysis of the relationship between these proto-forms and their reflexes is somewhat precarious. Their precise history is likely unreconstructable, and cannot be known for sure. Blust is not ignorant of this:

Each of the case studies [...] in the present article has been examined with this possibility in mind, and for each of them a theoretically orthodox interpretation was rejected as being a product of theory-driven speculation rather than of observation-based analysis. In some cases, my skepticism may have been misguided, but it was intended to “put the brakes” on rapid dismissal of such cases through superficial speculations about possible unseen developments, rather than seriously considering the possibility that in some cases the actuation of change is a product of conscious choice. (Blust 2017, 365).

Leaving aside for now Blust’s reference to “conscious choice”, it is important to both be open minded to Blust’s unorthodox analysis, while also keeping in mind his own admission of potential “misguided skepticism” throughout the subsequent discussion of his data.

2.1 Final *a Mutation

The first diachronic peculiarity that Blust draws attention to is final *a mutation, a phenomenon present in at least twenty-six daughter languages of Proto-Austronesian (PAN) in which PAN *a became a different vowel in word-final position (most frequently [o], but also [e],[o],[ə], [ɛ], [ʌ], or [i]). What Blust considers so peculiar about this specific change is its subject, and its conditioning environment. Blust holds that vowels are relatively stable during historical change cross-linguistically, and that [a] specifically is thought to be the most stable. Additionally, Blust holds that there is nothing about word final environments that would cause [a] to shift, nor could this change be ascribed to assimilation triggered by high vowels in other syllables. The fact, then, that it underwent a conditioned change in twenty-six languages is absolutely unexpected, and may require an explanation beyond normal phonetic processes.
Is this change, however, as strange as Blust makes it out to be? Appealing to the perceptual parsing bias error in Garrett & Johnson’s theory of sound change, certain vowels, when misheard, are more likely to be misheard in specific ways. In an English language study, Peterson & Blarney (1952) found experimentally that the two most common mishearings of [a] were [ɔ] and [ʌ], accounting respectively for ~10% and ~2% of vowel identifications of a segment containing [a]. As both of these forms are present in the reflexes of *a, this could seem promising for an account of final *a mutation driven by perceptual parsing bias. Even with this support, though, there are still two issues with this analysis: the first is that while [ɔ] and [ʌ] are present in the set of reflexes, the reflex with the most frequency by far is [ɔ], and the second is that even with an established bias for [a]’s misperception, this explanation alone cannot account for the change’s word final conditioning environment. The issue of [ɔ]’s predominance as a reflex, however, can be accounted for in two ways. Either one could propose an intermittent step through which [a] > [ɔ] > [ɔ], or one could suggest that [a] > [ɔ] directly, given that *o would have been the nearest vowel to [ɔ] in the PAN vowel space, and that a sound realized as [ɔ] would then most likely merge with /ɔ/ phonemically.

This leaves, however, the issue of a word-final conditioning environment. While this indeed does seem difficult to account for, possible interaction with the PAN stress system might shed some light on why this change could have occurred in the position that it did. This is addressed in the following section.

2.2 Final Syllable *u Fronting

Final syllable *u fronting refers to the fronting of Proto-Malayo-Polynesian (PMP) *u to [i] when in the final syllable. This occurred in several languages in eastern Indonesia. Again, this
change is peculiar for two reasons: a shift from [u] to [i] without the presence of a fronting phone is unexpected, and the fact that its conditioning environment is simply within the final syllable leads to further confusion.

Unlike in final *a mutation, an appeal to a perceptual parsing error bias is not empirically founded. In Peterson & Blarney (1952) not a single utterance of [u] was misheard as [i]. With that being said, however, what final syllable *u fronting and final *a mutation do have in common is a shared conditioning environment in word final syllables. Given this connection, including another change in which *u lowers to [o] in penultimate syllables preceding high vowels in Sarawak, and another in which low vowels raise in penultimate syllables in Oceanic, one might be able to posit some sort of syllable-sensitive phonetic pressure in ancestral forms of Austronesian that has either been lost, or has not yet been described in modern varieties. Such an explanation is not entirely out of the question. While the stress system of PAN has not been fully reconstructed, data from Formosan and Philippine languages do provide relevant information as to how stress may have been realized in PAN (Wolff 1993). Crucially, it seems that stress came in the final and penultimate syllables, occurring variably between them. A lack of stress has associations with changes in vowel quality, particularly devoicing or reduction (Ohala 1983, 203), processes that may have been made more likely due to the adjacency of a stressed syllable. It could then be possible that devoicing or reduction of unstressed vowels could have altered their quality in a way that was later phonologized into the languages in the relevant final or penultimate syllable positions. In the case of final *a mutation, the lowering of *u in Sarawak, and raising of low vowels in Oceanic, this could provide some explanation as to why these vowels changed. Given that vowel reduction often results in the loss of sonority and hence
centralisation of vowels, this would explain why in these instances low vowels undergo raising, and high vowels undergo lowering.

In the case of final syllable *u fronting, however, the situation is more complicated. Seeing that this change involves fronting rather than a lowering towards centralization, this change is inconsistent with the type of reduction proposed above. While this is certainly problematic, it should be noted that high vowels are already much more likely to undergo devoicing due to a reduction in transglottal pressure as compared to low vowels - the likelihood of devoicing would then only be compounded by a lack of stress (Ohala 1983). As such, it is possible that both *u and *i in this unstressed final syllable environment merged due their shared lack of voicing, becoming phonemically reinterpreted as /i/.

Regardless of the specific course of events that lead to these changes, the fact that they all seem to run parallel to the PAN stress system suggests that these changes are indeed motivated by factors that are fundamentally linguistic.

2.3 *a > i / q→q in Kavalan

In this change, reflexes of the PAN *a when adjacent to *q map to [i] in Kavalan, a language of Taiwan. This is perhaps the most problematic entry in Blust’s data, though it also might have the most convincing phonetic explanation. Given that this is a change from PAN to modern Kavalan without a similar change attested in any related languages that could suggest an earlier ancestral moment in which this change occurred, the estimated distance between this proto-form and its reflex could span five millenia (Gray et al. 2009). This amount of time could certainly allow for any number of intermediate changes in between the proto-form and reflex, and it is not

1 *q itself was lost from the phoneme inventory, and does not have any reflexes in Kavalan, though /q/ later re-emerged after a sound change lowered /k/ (Li & and Tsuchida 2006).
at all out of the question to imagine a completely phonetically motivated path leading to the current mapping. Such a path, however, may not even be necessary - a reasonable phonetic explanation for this mapping can be made with only one step.

What makes this sound change seem phonetically opaque is the unexpected raising of a low vowel when adjacent to a uvular stop, a phone which typically has a lowering effect on adjacent phones. However, the unexpected raising of a vowel in a typically low environment is precisely the sort of dissimilating change that can be explained by hypercorrection, particularly as an anticipatory rather than perseverative change. Given that listeners might correct the lowering influence of a uvular stop when adjacent to high vowels, it is in keeping with Ohala’s paradigm of hypercorrection to expect that low vowels in typically lowering-inducing environments might be raised by listeners who are erroneously “accounting” for a lowering effect. Another similar possibility is to understand this segment as ambiguous, and analyze it through Blevin’s CHANCE.

2.4 \( *b > k / V_V \) in Berawan

One of the most unintuitive changes presented by Blust is the intervocalic backing and devoicing of voiced labial stops in Berawan from Proto-Northern Sarawak (PNS). Seeing devoicing as well as a radical shift in place of articulation to be both conditioned intervocically is indeed puzzling. Blust, perhaps unexpectedly, proposes that this change occurred in two stages - an initial backing of \( *b > *g / V_V \), and then a devoicing of \( *g > k / V_V \).

While such a change does seem hard to account for, similar phenomena have been discussed in the literature surrounding phonetic change. While there are gaps in how this precise change can be explained, aspects of it have been addressed in appeals to the aerodynamics of
speech production (thus invoking aerodynamic bias factors as discussed by Garrett & Johnson).

First of all, stops are relatively more difficult to voice than other phones due to a greater air pressure above the glottis than is typically present in the production of other manners of articulation. As voicing requires a greater pressure below the glottis than above it, the smaller the pressure differential, the more difficult it becomes to voice a phone. As such, stops are already more predisposed to lose voicing than are other manners of articulation (Ohala 1983).

Additionally, certain stops are more likely to lose voicing than others - namely velars (Ohala 1993). This is because stops produced in the front of the oral cavity allow for a greater volume of air, thus lessening air pressure above the glottis. Velars, however, with their place of articulation in the back of the oral cavity, do not provide the same amount of space for air to fill, resulting in a higher air pressure above the glottis and consequently greater difficulty in maintaining voicing. While this explanation is complicated by the intervocalic environment in which this change occurs, the fact of aerodynamic bias factors in velar stop voicing does help in sorting out how such a change can be accounted for phonetically. Further, one possible way to account for the intervocally conditioned change is to once again propose hypercorrection. Intervocalic devoicing certainly fits the definition of a dissimilation and is thus consistent with an appeal to hypercorrection, though this explanation for this particular sound change does seem somewhat lacking.

The next consideration is how a labial stop could become velar, a change which occurred both in Berawan and in Palñawan dialect of Ayatal, in which word final labials became velar (*p > k / _#). This is more difficult to explain. Somewhat similar changes have been commonly seen cross-linguistically, though not in the same direction. Labialized voiceless velar stops becoming
voiceless labial stops (*kʷ > p) is one phenomenon which characterized the phonological shifts from Proto-Indo-European to Ancient Greek. This change may be related to a perceptual parsing bias, or Blevins’ change, in which perceptual similarities are the guiding force in sound change. Proposing such a change in the opposite direction for similar reasons is not absurd, but is also not much more than speculation, particularly when one considers that a labialized stage of the velar is unattested.

An alternative analysis to this sound change is proposed in Begus (2018). This approach looks at this specific change within the greater context of phonological changes in Berawan and proposes a phonetically motivated sequence of changes to account for the data. This sequence is as follows:

1. *b > *β / V_V
2. *β > *∅ /
3. *∅ > *x /
4. *x > k /

Here, *b lenites to a fricative in intervocalic environments, but remains a stop word initially. Following this, the bilabial fricative undergoes devoicing, though the apparent intervocalic environment is perhaps deceptive. Begus notes that the conditions for frication and voicing are “diametrically opposed” - one requiring high oral pressure and the other low oral pressure. Reminiscent of the discussion of the phonetics of stop voicing above, fricative voicing involves similar issues, though their production is perhaps further complicated by their longer duration. As such, unconditioned fricative devoicing, according to Begus, is a common and expected change - the fact that it appears to be intervocally conditioned in Berawan is solely because the only fricatives happened to be intervocalic. Following this, there is a change from *∅ > *x
which could be linked to perceptual similarity (and thus attributable to change or perceptual parsing bias), an explanation for which there exists "some evidence" (Begus 128). Finally, a change from *x > k, a fortition, which Begus claims is motivated by a "reduction of articulatory precision" (Begus 129) finishes the sequence and established what Begus calls a "Blurring Chain", a sequence of sound changes in which the last step "blurs" phonetic motivation between proto-form and reflex - precisely the sort of sequence required by a mainstream interpretation of historical linguistics in order to explain Blust's data.

3 Consequences, Alternatives, and Concluding Remarks

One of the fundamental issues coming out of Blust's data and any potential analysis of it is the question of to what extent one can be comfortable with theory-driven explanations that lack direct evidence, and whether looking towards explanations outside the realm of phonetic motivation provides a superior analysis. With certainty, Blust's data complicates any notion of universal phonetically motivated internal change. While feasible phonetic explanations can be provided for some examples, others are still quite mysterious, and no provided explanation is completely tight. Though Blust's counterexamples are not absolutely damning to a theory of phonetically motivated change (it is important to keep in mind the generally vast timescales involved in Blust's mappings and the potential for intermediate stages within these mappings), they do make the position less tenable. Still, it should not be ignored how powerful phonetic explanations are for sound change. Phonetic explanations are intuitive, and are largely applicable to the majority of attested sound change data. Because of this, what Blust refers to as "theory-driven speculation" may not so much be the grasping-at-straws, post-hoc band aid that he makes it out to be, but rather it could be seen as the application of Occam's Razor in a
situation where the crucial data might be irrecoverable. Indeed, it would seem that theory exists precisely for this reason - to allow one to fill in gaps in the data in a way that is consistent with what is verifiably known.

Still, what might a few inconsistent data points mean for the theory at large, assuming that they truly are not phonetically motivated (a conclusion that as of yet is not convincingly supported)? The consequences may not be so severe. Given that some amount of uncertainty is inherent in human behavior, allowing for a probabilistic theory that is successful in the vast majority of cases, while allowing for a small amount of inconsistency and variability born out of the unpredictability of human beings seems absolutely reasonable, and is a model followed throughout the social sciences.

Such a model, however, is not without its detractors. Blust’s data withstanding, there are compelling reasons to be dubious of any number of linguistic explanations suggested for diachronic change. Several concerns regarding the explanatory power of phonetic and linguistic explanations for sound change were presented in Lass (1980), including a severe critique of probabilistic explanations, which, to Lass, are not explanations at all. In his discussion of explanations rooted in naturalness, Lass brings up two crucial issues involved in linguistic accounts of why sound change occurs - the multiple and null strategy problems. The multiple strategy problem refers to the fact that for two instances of sound change with identical initial conditions, two very different outcomes can occur, while the null strategy problem refers to the fact that for two identical environments, a sound change might occur in one while another undergoes no change at all. Lass’ fundamental issue then with phonetic motivation as an explanatory model, is that it lacks sufficient falsifiability. According to Lass, any single instance
of inconsistency within this framework is not enough to scrap theory, but rather, theory is only abandoned once the number of incompatible data points passes an arbitrary threshold. As such, to Lass, “explanations” of sound change lack any true, empirically based explanatory power. While many might consider this lack of absolute empiricism acceptable within a discipline like linguistics, a field that is bound to human behavior, this state of affairs becomes particularly problematic when one is attempting to address diachronic changes that lack direct attestation. If the model of explanation itself brings with it a relatively high level of unpredictability, then it follows why Blust might be so resentful of a perceived over-reliance on theory-driven linguistic explanations when these explanations are far from having perfect consistency.

Given this, it is salient to bring up the final question posed by Blust at the end of his paper - must sound change be linguistically motivated? Though linguistic explanations can be powerful and convincing, there is no reason to suppose that they are always necessary, particularly when the data seemingly cannot support them. Rather than maintain what are at times admittedly somewhat convoluted phonetic explanations for his data, Blust appeals to the possibility of “conscious choice” playing a role in regular sound change. This is an interesting suggestion, but unfortunately Blust does little to elaborate on or support this position, saying no more than the following:

[Despite cases of phonetically motivated change,] conscious choice plays a role in various other sound changes, and in these changes speakers may make use of an innate knowledge of natural classes to produce phonological innovations of a purely conventional nature (Blust 2017, 367).

Exactly how conscious use of this “innate knowledge” is manifested or proven is unclear. As such, with the given information, explanations that appeal to “conscious choice” seem no more empirically founded than those assuming intermediate stages of phonetic changes. Still, though
perhaps posed problematically by Blust, an investigation of what a “conscious choice” model for sound change could look like is worth considering.

3.1 Conscious Choice

It is relevant to note that the idea of “choice” has already been incorporated, at least to some extent, into a pre-existing model of phonetically motivated change. This is the third mechanism of Blevins’ sound change typology, conveniently named CHOICE. In CHOICE, listeners are presented with several acceptable options for the pronunciation of a given segment, often conditioned by social factors or register. From this pool of acceptable pronunciations, the listener “chooses” (unconsciously) to internalize an exemplar form which does not match with the intended form of the speaker. This difference in the listener’s internalized form from the speaker’s intended form is what then constitutes a sound change. What differentiates CHOICE from the two other mechanisms in Blevins’ sound change typology, as well as those discussed in other theories of phonetically motivated change, is that it already presupposes in some sense the actuation of a change. In CHOICE, it is already assumed that speaker variation (due to phonetic factors) has altered the phonetic signal, and that perceptibly different forms are being uttered variably for the same underlying form - something perhaps akin to lexical diffusion. Here, the choice that takes place is not of an active desire to change how a segment is realized, but rather a choice between pre-existing forms determined by register or frequency.

Indeed, most discussions of the involvement of conscious or semiconscious choice in sound change apply to a period that follows the existence of an altered phonetic signal. This is rather different form the sort of sound change Blust seems to be proposing. It is not so difficult to imagine how an individual’s choice could play a role in the propagation of a sound change; it is
much more difficult to imagine how an individual’s choice could play a role in the systematic actuation of a sound change. A scenario in which phonetically motivated variation leads one to chose to lenite stops to fricatives intervocalically because one has heard this done around them is much easier to follow than a scenario in which an individual consciously decides to systematically make this change without precedent. It seems, however, that this is precisely the sort of conscious action that Blust is suggesting has occurred throughout his data.

Instances of speakers deliberately changing their language in such a way that it could complicate internal reconstruction or the comparative method is perhaps surprisingly not unheard of. Systematic grammatical changes that originate in a desire to separate one’s own group from another have been documented cross-linguistically. For example, the systematic reversal of gender markers in Usai (a language of New Guinea) and a pattern of metathesis in Quechua (though not a systematically coherent one) seem to have occurred for precisely this reason (Thomason 1999). In a discussion of deliberate change, Thomason gives the following account for how and when such changes can take place and have a lasting effect:

First, speakers’ choices can indeed lead to drastic linguistic changes. Second, these changes only rarely have a permanent effect on the speech of an entire community; and where they do have a permanent effect, it is because of particular social circumstances. First, the speech community must be small. But in addition, there must be other contributing social factors, though not all of them can be identified on the basis of currently available information. One common factor is very widespread multilingualism, with or without socioeconomic dominance by one group in the contact situation, so that the tension between an other-directed world view and a self-directed worldview may come into play. Another potential contributing factor, probably less common, is the deliberate actions of language standardizers. A third is the emergence of a new ethnic group that seeks a language to symbolize its new identity (Thomason 1999, 39).

The account given here lends credence to Blust’s proposition of conscious choice for some of his data points, but is inconsistent with others. Considering the intervocalic stop backing and
devoicing from Berawan, Thomason’s criteria for deliberate change is compelling. The speech community of Berawan is indeed small, with only 1500 speakers (Ethnologue). It is also spoken on the island of Borneo, a region with a relatively high index of linguistic diversity. Moreover, the Berawan exist in a peculiar way within the greater ethno-cultural context of coastal Sarawak. Though they claim to be a part of the Kenyah ethnic group, their customs and language are decidedly different from Kenyah. Rather, the Berawan show a greater linguistic affinity with neighbors in the Lower Baram group, groups with whom they also share a funerary practice called *mdang*, which is not practiced by the Kenyah (Metcalf 1976). In addition to this, a subset of the Berawan are quite possibly the only practitioners of *aded luna*, a religious practice that predates the introduction of Abrahamic religions to the region. All of this information together paints an interesting picture relating to Thomason’s criteria. It seems that the Berawan indeed do exhibit cultural practices that would separate them from groups in related, adjacent speech communities, and that they have actively separated themselves in adopting the Kenyah demonym. Given this, it is not entirely impossible to imagine a situation in which speakers of Berawan (or a subset of them) deliberately altered their speech in order to distance themselves from another group. However, there is still no direct evidence of this occurring, and cultural separation itself can be quite powerful in propagating any kind of linguistic change, regardless of how it originates. Finally, it should also be noted that Proto-Northern-Sarawak, the proto-language from which this sound change has been constructed, was spoken 3000 years before Berawan - a timescale that most definitely allows for any number of sound changes, including the four suggested in Begus’ blurring chain.
Further, several of Blust's examples like final syllable *u fronting or final *a mutation are quite inconsistent with Thomason's criteria, given that they took place across an expansive geographical area among a diverse set of speech communities. While indeed some of the speech communities involved may have been small enough and had the proper cultural circumstances for deliberate language change to take place, such an explanation does not make sense within the larger image that these sound changes create; that is, one which strongly resembles drift brought on by shared inherited phonetic pressure. Even if the exact character of this pressure is difficult to conceptualize, this explanation seems far more coherent than a conscious choice model.

It also seems salient to again refer to the null strategy problem. Thomason cautions the following:

It must be emphasized, however, that no contributing factors, no matter how powerful they are in some contexts, will permit us to predict when speakers' choices will produce major changes in a language: contributing social factors are necessary conditions for the kinds of changes we're talking about, but not sufficient conditions. Even where small groups live as close neighbors, with very extensive mutual multilingualism, we don't always find widespread structural convergence; whether it occurs or not depends on cultural factors that are likely to remain permanently beyond our predictive grasp (Thomason 1999, 39).

From this, one sees that a model of consciously induced language change suffers from one of the same fundamental issues involved in phonetically motivated models of sound change - neither has strong predictive validity, and the information needed to make explicit predictions is probably beyond the scope of human capacity. Further, it is uncontroversial to say that the phonetic factors that pattern with sound change are far better understood than the cultural factors associated with deliberate language change. Given this, it is difficult to imagine why one would prefer a non-phonetic model in a situation where the only evidence to suggest that a far rarer, less understood phenomenon may have occurred is the fact that a proto-form/reflex mapping appears
opaque. Concluding deliberate language change is then further complicated by the fact that these "opaque" mappings, as I have shown, are not impossibly incompatible with the application of phonetic theory, and that systematic conditioned phonological changes have yet to be documented as a subset of possible deliberate language changes. While further developments in linguistic anthropology and sociolinguistics might warrant a reevaluation of this position, appeals to conscious choice will almost always be dispreferred to linguistic explanations unless sufficient evidence is provided. In the data provided by Blust, there is very little discussion or even conjecture about the cultural contexts in which his sound changes took place, and as such, there is little reason to assume that the cultural factors necessary to bring on deliberate language change were present.

This does not mean, however, that this knowledge, or other relevant pieces of information, are unobtainable, or that their pursuit is a dead end. Conscious choice in language change is a fascinating and poorly understood phenomenon that deserves greater scrutiny, and, as I have shown, there are compelling reasons to consider that deliberate language change may have been a possibility in the change seen in Berawan. That being said, this sort of language change is far less common than phonetically motivated change, and in situations where so little intermediate information is known, it seems strange, all else being equal, to prefer non-phonetic explanations. Skepticism is incredibly valuable, and it is crucial for the development of the field that alternative explanations are considered whenever the data is inconsistent with established theory. However, this does not mean that unexpected gaps in the data will always lend themselves to unorthodox explanations. Ultimately, Blust’s data, though mysterious, is not inconsistent with phonetic motivation, particularly when one considers how drastically languages
can change in the timescales being considered. Rather, it may be preferable to see the mystery of Blust’s data in a different light. Instead of searching for an explanation beyond language, one should be reminded of how little is still known about the psycholinguistics of sound change and phonetic perception. Perceptual parsing bias has been proposed as an explanation for several of the sound changes discussed in this paper, though Garrett & Johnson themselves maintain that this phenomenon is not well understood. Despite this, biases born from perceptual confusion have been experimentally identified, as in the experiment by Peterson & Blarney referenced in the section on final *a mutation. Further experimentation in this vain cross-linguistically may lead to insights useful in addressing cases of unexpected sound change, as well as lead towards a fuller understanding of the perception and misperception of phonetic signals.

**References**


29 19-43.